

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 08:36:32 ; Search time 54 Seconds

(without alignments)
1072.634 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047
Sequence: 1 MSKKKGLAEKRTMELP.....FEENKIDRTFGPEDFDYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1047	100.0	205	7	ADD84537
2	1047	100.0	219	6	ABP75541
3	975	93.1	190	4	AAM40043
4	975	93.1	190	7	ADD84547
5	945.5	90.3	190	7	ADD84545
6	606	57.9	122	4	AAM41829
7	591	56.4	122	7	ADD84543
8	584	55.8	119	7	ADD84541
9	465	44.4	126	7	ADD84539
10	329.5	31.5	182	4	ABG12341
11	287	27.4	79	4	AAG74669
12	134.5	12.8	1258	6	ABP98829
13	134.5	12.8	1266	6	ABR57415
14	134.5	12.8	1302	4	AAM38659
15	125.5	12.0	266	7	ADE15628
16	121.5	11.6	1881	5	ABP73809
17	120.5	11.5	1798	4	ABB66937
18	117	11.2	795	5	ABB77430
19	117	11.2	976	4	RAG65581
20	117	11.2	976	6	ABP74709
21	117	11.2	976	6	ABP74709
22	117	11.2	976	6	ABP74709
23	116.5	11.1	1328	6	ABO14658
24	116.5	11.1	1329	7	ADE15650
25	116.5	11.1	2383	5	ABG65631

26	116.5	11.1	2442	3	RAY77575
27	115.5	11.0	880	4	AAB96332
28	114	10.9	284	7	ADB67124
29	113.5	10.8	284	7	ADB67131
30	113	10.8	484	4	AAM78985
31	113	10.8	533	4	AAM79969
32	113	10.8	1242	7	ADB70309
33	112	10.7	284	7	ADB67112
34	111.5	10.6	1327	5	ABBS7163
35	110.5	10.6	284	6	ABO14687
36	110	10.5	284	7	ADB67118
37	110	10.5	284	7	ADB67115
38	110	10.5	972	6	ABJ19394
39	109.5	10.5	168	4	AAE12789
40	109.5	10.5	168	7	ABO23548
41	109.5	10.5	818	7	ADB64392
42	109	10.4	284	7	ADB67120
43	108.5	10.4	284	4	AAM78512
44	108.5	10.4	860	7	ADC38517
45	107.5	10.3	1185	6	ABU19878
46	107	10.2	284	7	ADB67132
47	107	10.2	284	3	RAY92334
48	107	10.2	284	4	AAG66545
49	107	10.2	284	5	ABBS90770
50	107	10.2	284	5	ABBS7037
51	107	10.2	284	6	ABO14685
52	107	10.2	284	6	ABU54477
53	107	10.2	284	6	ABR92157
54	107	10.2	284	7	ADB67109
55	107	10.2	284	7	ADB67117
56	107	10.2	284	7	ADB67119
57	107	10.2	284	7	ADE63267
58	107	10.2	284	7	ADD48685
59	107	10.2	284	7	ADE57458
60	107	10.2	284	7	ADE57462
61	107	10.2	284	7	ADE60981
62	107	10.2	284	7	ADD47178
63	107	10.2	284	7	ADD47180
64	107	10.2	284	7	ADE57464
65	107	10.2	284	7	ADD48687
66	107	10.2	284	7	ADE57460
67	107	10.2	287	7	ADE15642
68	107	10.2	888	5	AAU83013
69	106.5	10.2	284	7	ADB67110
70	106	10.1	243	5	ADB65013
71	106	10.1	410	5	ABP43547
72	106	10.1	534	2	AAW46823
73	106	10.1	534	2	AAW98948
74	106	10.1	1558	3	AAE18324
75	106	10.1	1786	2	AAW24790
76	106	10.1	1787	5	AAU98699
77	105	10.0	245	7	ADB67134
78	105	10.0	251	7	ADB67133
79	105	10.0	284	7	ADB67129
80	105	10.0	975	6	ADA54467
81	105	10.0	1846	6	ABR39833
82	105	10.0	1912	7	ADC28223
83	104.5	10.0	419	4	ABP96505
84	104.5	10.0	1177	3	AAE21231
85	104.5	10.0	1404	6	ABP55393
86	104.5	10.0	248	5	ABP97266
87	104	9.9	1939	7	ADB47857
88	104	9.9	1939	7	ADB47861
89	104	9.9	1939	7	ADB47861
90	104	9.9	1939	7	ADB47861
91	103.5	9.9	1122	6	ABU70418
92	103.5	9.9	1124	7	ADC37285
93	103.5	9.9	1193	4	AAE25602
94	103.5	9.9	1193	5	ABG61824
95	103.5	9.9	1194	4	AAU32407
96	103.5	9.9	1197	4	AAU93006
97	103.5	9.9	1384	6	ABP55413
98	103.5	9.9	1384	6	ABP55413

RAY77575	Human cyt
ABP96332	Putative
ADB67124	Tropomyos
ADB67131	Tropomyos
AAM78985	Human pro
AAM79969	Human pro
ADB70309	C. neofor
ADB67112	Tropomyos
ABBS7163	Mouse isc
ABO14687	Novel hum
ADB67118	Tropomyos
ADB67115	Tropomyos
ABJ19394	Human int
AAE12789	Chlamydia
ABO23548	Chlamydia
ADB64392	Human pro
ADB67120	Tropomyos
AAM78512	Human pro
ADC38517	Human AML
ABU19878	Protein e
ADB67132	Tropomyos
RAY92334	Human alp
AAG66545	Human int
ABBS90770	Human int
ABBS7037	Mouse isc
ABO14685	Novel hum
ABU54477	Human hum
ABR92157	Human cer
ADB67109	Tropomyos
ADB67117	Tropomyos
ADB67119	Tropomyos
ADE63267	Rat Prote
ADD48685	Rat Prote
ADE57458	Rat Prote
ADE57462	Rat Prote
ADE60981	Rat Prote
ADD47178	Rat Prote
ADD47180	Human Pro
ADE57464	Human Pro
ADD48687	Human Pro
ADE57460	Human Pro
ADE15642	Human str
AAU83013	Human hom
ADB67110	Tropomyos
ADB65013	Human pro
ABP43547	Inner cen
AAW46823	Amino aci
AAW98948	Streptoco
AAE18324	Plasmodiu
AAW24790	P. falcip
AAU98699	Plasmodiu
ADB67134	Tropomyos
ADB67133	Tropomyos
ADB67129	Tropomyos
ADA54467	Human pro
ABR39833	Human SCA
ADC28223	Human NOV
ABP96505	Putative
AAE21231	Tomato le
ABP55393	Putative
ABP97266	Human MDD
ADB47857	Human pro
ADB47861	Human pro
ADB47861	Human pro
ADB47861	Human pro
ABU70418	Nuclear f
ADC37285	Nuclear f
AAE25602	Human pro
ABG61824	Prostate
AAU32407	Novel hum
AAU93006	Human pro
ABP55413	Human MDD

99 103.5 9.9 1591 7 ADE40160 Human NOV
100 103.5 9.9 1907 5 ABB82127 Human TAN

ALIGNMENTS

RESULT 1
ADD84537
ID ADD84537 standard; protein; 205 AA.

XX AC ADD84537;
XX DT 29-JAN-2004 (first entry)
XX DE 121P1F1 protein.
XX KW 121P1F1; 121P1F1 modulation; human; chromosome 4q; cytostatic;
XX KW gene therapy; vaccine; cancer; immune response; immunisation.
XX OS Homo sapiens.
XX PN WO200295009-A2.
XX PD 28-NOV-2002.
XX PF 28-FEB-2002; 2002WO-US006242.
XX PR 05-MAR-2001; 2001US-00799250.
XX PA (AGEN-) AGENSYS INC.
XX PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
XX PI Jakobovits A;
XX DR WPI; 2003-156757/15.
XX DR N-PSDB; ADD84536.
XX PT Composition comprising a substance that modulates the status of 121P1F1,
XX PT useful in diagnosing, preventing, prognosticating or treating patients
XX PT with cancer that expresses 121P1F1, such as breast, colon, ovarian or
XX PT lung cancer.
XX PS Claim 19; Fig 2A; 285pp; English.
XX CC The present invention describes a composition (I) comprising a substance
XX CC that modulates the status of 121P1F1 (gene and encoded protein), or a
XX CC molecule that is modulated by 121P1F1, where the status of a cell that
XX CC expresses 121P1F1 is modulated. The human 121P1F1 gene maps to chromosome
XX CC 4q. (I) has cytostatic activity, and can be used in gene therapy, and in
XX CC vaccines. The composition (I) can be used for diagnosing, preventing,
XX CC prognosticating or treating patients with cancer that expresses 121P1F1,
XX CC such as breast, colon, ovarian or lung cancer. The 121P1F1 gene or its
XX CC fragment can be used to elicit a humoral or cellular immune response.
XX CC 121P1F1 antibodies can be used in active or passive immunisation. 121P1F1
XX CC polynucleotides are useful as probes and primers for the amplification or
XX CC detection of 121P1F1 genes, as coding sequences for directing the
XX CC expression of 121P1F1 polypeptides, or as tools for modulating or
XX CC inhibiting the expression of 121P1F1 genes. The present sequence is used
XX CC in the exemplification of the present invention.
XX SQ Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 7; Length 205;
Best Local Similarity 100.0%; Pred. NO. 1.5e-86;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKGLSAREKTRMWEIFSETKDFQLKLEKAPKEKGTAMSVKVELQSLVDDGMY 60
DB 1 MSKKGLSAREKTRMWEIFSETKDFQLKLEKAPKEKGTAMSVKVELQSLVDDGMY 60
QY 61 DCRIGTSNYWAPPSKALHARKHKLVEYLSQSGSQHASLQKSIKAKIGRCETEER 120

Db 61 DCRIGTSNYWAPPSKALHARKHKLVEYLSQSGSQHASLQKSIKAKIGRCETEER 120
QY 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 180
DB 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 180
QY 181 KRKGFEEKNKIDRTFGIPEDFDYID 205
DB 181 KRKGFEEKNKIDRTFGIPEDFDYID 205

RESULT 2
ABP75541

ID ABP75541 standard; protein; 219 AA.

XX AC ABP75541;
XX DT 10-FEB-2003 (first entry)
XX DE Human secretory polypeptide SPTM SEQ ID NO 725.
XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;
XX KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;
XX KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;
XX KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;
XX KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;
XX KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;
XX KW antipsoriatic; anti-anaemic; anti-HIV; human immunodeficiency virus;
XX KW secretory polynucleotide; secretory protein.
XX OS Homo sapiens.
XX PN WO200293876-A2.
XX PD 24-OCT-2002.
XX PF 27-MAR-2002; 2002WO-US009921.
XX PR 29-MAR-2001; 2001US-0280067P.
XX PR 29-MAR-2001; 2001US-0280068P.
XX PR 16-MAY-2001; 2001US-0291280P.
XX PR 17-MAY-2001; 2001US-0291829P.
XX PR 17-MAY-2001; 2001US-0291849P.
XX PR 19-JUN-2001; 2001US-0299428P.
XX PR 20-JUN-2001; 2001US-0299776P.
XX PR 20-JUN-2001; 2001US-0300001P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Dafo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld V, Gerstin EH;
XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX PI Flores V, Marwaha R, Lo A, Lan XY, Urashka WE;
XX WPI; 2003-075543/07.
XX DR N-PSDB; ABZ95987.
XX PT New human secretory proteins and polynucleotides, useful for diagnosing,
XX PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),
XX PT neurological disorders (e.g. Alzheimer's), or cell proliferations or
XX PT cancers.
XX PS Claim 27; SEQ ID NO 725; 458pp + Sequence Listing; English.
XX CC The invention relates to a secretory polynucleotide (designated sptm)
XX CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a
XX CC naturally occurring polynucleotide sequence at least 90 % identical to
XX CC the polynucleotide sequence, a polynucleotide complementary to them or an
XX CC RNA equivalent of them. The polypeptide or polynucleotide are useful for
XX CC treating, preventing or diagnosing a disease or condition associated with
XX CC the expression of functional SPTM. These are particularly useful for
XX CC diagnosing, treating or preventing autoimmune/inflammatory disorders

(e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). The present sequence is one of the SPM proteins of the invention (ABP75384-ABP75962). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 219 AA;

Query Match 100.0%; Score 1047; DB 6; Length 219;
 Best Local Similarity 100.0%; Pred. No. 1.6e-88;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKGLSAEKRTRWEIFSTKDVFLKDLKIAPEKIGITAMSVKEVLQSLVDDGMV 60
 DB 15 MSKKGLSAEKRTRWEIFSTKDVFLKDLKIAPEKIGITAMSVKEVLQSLVDDGMV 74
 QY 61 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSQHASLQKSIKAKIGRCETEER 120
 DB 75 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSQHASLQKSIKAKIGRCETEER 134
 QY 121 TLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIPAKSWA 180
 DB 135 TLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIPAKSWA 194
 QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
 DB 195 KRKFGFEENKIDRTFGIPEDFDYID 219

RESULT 3
 AAM40043
 ID AAM40043 standard; protein; 190 AA.
 AC AAM40043;
 XX
 XX 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 3188.
 DE
 DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
 XX
 XX Homo sapiens.
 XX WO200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 XX 21-JAN-2000; 2000US-0048725.
 XX 25-APR-2000; 2000US-00552317.
 XX 20-JUN-2000; 2000US-00598042.
 XX 19-JUL-2000; 2000US-00620312.
 XX 03-AUG-2000; 2000US-00653450.
 XX 14-SEP-2000; 2000US-00662191.
 XX 19-OCT-2000; 2000US-00693036.
 XX 23-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou F, Goodrich R, Drmanac RI;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI59199.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 4; SEQ ID NO 3188; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAI42213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX Sequence 190 AA;
 SQ

Query Match 93.1%; Score 975; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 6e-82;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSETKDVFLKDLKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAF 75
 DB 1 MMEIFSETKDVFLKDLKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAF 60
 QY 76 SKALHARKHKLKLEVLSESGSQHASLQKSIKAKIGRCETEERTRAKELSLRDQRE 135
 DB 61 SKALHARKHKLKLEVLSESGSQHASLQKSIKAKIGRCETEERTRAKELSLRDQRE 120
 QY 136 QLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIPAKSWAKKFGFEENKIDRTF 195
 DB 121 QLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIPAKSWAKKFGFEENKIDRTF 180
 QY 196 GIPEDFDYID 205
 DB 181 GIPEDFDYID 190

RESULT 4
 ADD84547
 ID ADD84547 standard; protein; 190 AA.
 XX ADD84547;
 XX 29-JAN-2004 (first entry)
 DT 121P1F1 variant 4 protein.
 DE
 DE 121P1F1; 121P1F1 modulation; human; chromosome 4q; cytostatic;
 KW gene therapy; vaccine; cancer; immune response; immunisation.
 XX Synthetic.
 XX Homo sapiens.
 XX WO200295009-A2.
 XX 28-NOV-2002.
 XX 28-FEB-2002; 2002WO-US006242.
 XX

```

PR 05-MAR-2001; 2001US-00799250.
PA (AGEN-) AGENSYS INC.
XX
XX Challita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
PI Jakobovits A;
XX
XX WPI; 2003-156757/15.
DR N-PSDB; ADD84546.
XX
XX Composition comprising a substance that modulates the status of 121P1F1,
PT useful in diagnosing, preventing, prognosticating or treating patients
PT with cancer that expresses 121P1F1, such as breast, colon, ovarian or
XX lung cancer.
XX
XX Claim 19; Fig 2F; 285pp; English.
XX
XX The present invention describes a composition (I) comprising a substance
CC that modulates the status of 121P1F1 (gene and encoded protein), or a
CC molecule that is modulated by 121P1F1, where the status of a cell that
CC expresses 121P1F1 is modulated. The human 121P1F1 gene maps to chromosome
CC 4q. (I) has cytostatic activity, and can be used in gene therapy, and in
CC vaccines. The composition (I) can be used for diagnosing, preventing,
CC prognosticating or treating patients with cancer that expresses 121P1F1,
CC such as breast, colon, ovarian or lung cancer. The 121P1F1 gene or its
CC fragment can be used to elicit a humoral or cellular immune response.
CC 121P1F1 antibodies can be used in active or passive immunisation. 121P1F1
CC polynucleotides are useful as probes and primers for the amplification or
CC detection of 121P1F1 genes, as coding sequences for directing the
CC expression of 121P1F1 polypeptides, or as tools for modulating or
CC inhibiting the expression of 121P1F1 genes. The present sequence is used
CC in the exemplification of the present invention.
XX
XX SQ Sequence 190 AA;
Query Match 93.1%; Score 975; DB 7; Length 190;
Best Local Similarity 100.0%; Pred. No. 66-82;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 MMEIFSETKDVFLKLEKIAPEKKEITAMSVKEVLSQSLVDGMDVCERIGTSNTYWAPP 75
DB 1 MMEIFSETKDVFLKLEKIAPEKKEITAMSVKEVLSQSLVDGMDVCERIGTSNTYWAPP 60
QY 76 SKALHARKHKLVEQLSEGSQKSHASLOKSIKAKIGRCETEERTRLAKELSSLRDQRE 135
DB 61 SKALHARKHKLVEQLSEGSQKSHASLOKSIKAKIGRCETEERTRLAKELSSLRDQRE 120
QY 136 QLKAEEVKYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSWAKRKFGEENKIDRTF 195
DB 121 QLKAEEVKYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSWAKRKFGEENKIDRTF 180
QY 196 GIPEDFDYID 205
DB 181 GIPEDFDYID 190
RESULT 5
ADD84545
ID ADD84545 standard; protein; 190 AA.
XX
XX AC ADD84545;
XX
XX 29-JAN-2004 (first entry)
XX
XX 121P1F1 variant 3 protein.
XX
XX 121P1F1; 121P1F1 modulation; human; chromosome 4q; cytostatic;
XX gene therapy; vaccine; cancer; immune response; immunisation.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO200295009-A2.

```

```

XX PD 28-NOV-2002.
XX
XX PF 28-FEB-2002; 2002WO-US006242.
XX
XX PR 05-MAR-2001; 2001US-00799250.
XX
XX PA (AGEN-) AGENSYS INC.
XX
XX PI Challita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
PI Jakobovits A;
XX
XX WPI; 2003-156757/15.
DR N-PSDB; ADD84544.
XX
XX Composition comprising a substance that modulates the status of 121P1F1,
PT useful in diagnosing, preventing, prognosticating or treating patients
PT with cancer that expresses 121P1F1, such as breast, colon, ovarian or
XX lung cancer.
XX
XX Claim 19; Fig 2E; 285pp; English.
XX
XX The present invention describes a composition (I) comprising a substance
CC that modulates the status of 121P1F1 (gene and encoded protein), or a
CC molecule that is modulated by 121P1F1, where the status of a cell that
CC expresses 121P1F1 is modulated. The human 121P1F1 gene maps to chromosome
CC 4q. (I) has cytostatic activity, and can be used in gene therapy, and in
CC vaccines. The composition (I) can be used for diagnosing, preventing,
CC prognosticating or treating patients with cancer that expresses 121P1F1,
CC such as breast, colon, ovarian or lung cancer. The 121P1F1 gene or its
CC fragment can be used to elicit a humoral or cellular immune response.
CC 121P1F1 antibodies can be used in active or passive immunisation. 121P1F1
CC polynucleotides are useful as probes and primers for the amplification or
CC detection of 121P1F1 genes, as coding sequences for directing the
CC expression of 121P1F1 polypeptides, or as tools for modulating or
CC inhibiting the expression of 121P1F1 genes. The present sequence is used
CC in the exemplification of the present invention.
XX
XX SQ Sequence 190 AA;
Query Match 90.3%; Score 945.5; DB 7; Length 190;
Best Local Similarity 92.2%; Pred. No. 3.2e-79;
Matches 189; Conservative 0; Mismatches 1; Indels 15; Gaps 1;
QY 1 MSKKKGLSAEKSTRMWEISETKDVFLKLEKIAPEKKEITAMSVKEVLSQSLVDGMDV 60
DB 1 MSKKKGLSAEKSTRMWEISETKDVFLKLEKIAPEKKEITAMSVKEVLSQSLVDGMDV 60
QY 61 DCERIGTSNTYWAPPSSKALHARKHKLVEQLSEGSQKSHASLOKSIKAKIGRCETEER 120
DB 61 DCERIGTSNTYWAPPSSKALHARKHKLVEQLSEGSQKSHASLOKSIKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEEVKYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSWA 180
DB 121 TRLAKELSSLRDQREQLKAEEVKYKDCDPQVVEIRQANKVAKAANRWTDNI FAIKSWA 165
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
DB 166 KRKFGFEENKIDRTFGIPEDFDYID 190
RESULT 6
AAM41829
ID AAM41829 standard; protein; 122 AA.
XX
XX AC AAM41829;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 6760.
XX
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KW

```


DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #12332.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX N-PSDB; AAS76528.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 42700; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (II) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities in
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AEG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 182 AA;
Query Match 31.5%; Score 329.5; DB 4; Length 182;
Best Local Similarity 61.3%; Pred. No. 2.6e-22;
Matches 73; Conservative 10; Mismatches 17; Indels 19; Gaps 2;
QY 47 VKEVLQSLVDGMDVCERIGTSNYWAFPSKALHAKHKLVELESQSGSKHASLQKS 106
Db 50 IKEILPKLI-----YRFNAKPI---KISAAFLQKLLSGSKHASLQKS 90
QY 107 IEKAKIGRCETETRTLAKELSLRDRBQLKAEVEKYKDCDPQVVEIRQANKYAKEA 165
Db 91 IEKAKIGRCETETRTLAKELSLRDRBQLKAEVEKYKDCDPQVVEIRPLVTLLKQA 149
RESULT 11
AAG74669
ID AAG74669 standard; protein; 79 AA.
XX

AC AAG74669;
XX 03-SEP-2001 (first entry)
XX Human colon cancer antigen protein SEQ ID NO:5433.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX Homo sapiens.
XX WO200122920-A2.
XX 05-APR-2001.
XX 28-SEP-2000; 2000WO-US026524.
XX 29-SEP-1999; 99US-0157137P.
XX 03-NOV-1999; 99US-0163280P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI: 2001-235357/24.
XX N-PSDB; AAH34074.
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX Claim 11; Page 7049-7050; 9803pp; English.
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
XX to 1052, 7921 and 7922
XX Sequence 79 AA;
Query Match 27.4%; Score 287; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 7.2e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 VEEIRQANKYAKEANRWTDNIPFAKSWAKRKFGFEENKIDRTFGIPEDFDYID 205
Db 26 VEEIRQANKYAKEANRWTDNIPFAKSWAKRKFGFEENKIDRTFGIPEDFDYID 79
RESULT 12
ABP98829
ID ABP98829 standard; protein; 1258 AA.
XX AC ABP98829;
XX 15-JUL-2003 (first entry)
XX Human structural and cytoskeletal associated protein #20.
XX Cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; antiangiinal;

KW neuroprotective; cerebroprotective; hypotensive; cardiatic; osteopathic;
 KW antiinflammatory; antiarthritic; virucide; gene therapy; human; stroke;
 KW structural and cytoskeleton-associated protein; SCAP; cancer; angina;
 KW atherosclerosis; epilepsy; Huntington's disease; hypertension;
 KW heart failure; osteoporosis; osteoarthritis.
 XX
 OS Homo sapiens.
 XX
 XX W02003031940-A2.
 XX
 XX 17-APR-2003.
 XX
 XX 10-OCT-2002; 2002WO-US032851.
 XX
 XX 12-OCT-2001; 2001US-032891P.
 XX
 XX 19-OCT-2001; 2001US-0360881P.
 XX
 XX 02-NOV-2001; 2001US-0343896P.
 XX
 XX 09-NOV-2001; 2001US-0346308P.
 XX
 XX 16-NOV-2001; 2001US-0332385P.
 XX
 XX 07-DEC-2001; 2001US-0340776P.
 XX
 XX 11-JAN-2002; 2002US-0347703P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD, Chang H;
 XX Chawla NK, Elliott VS, Emerling BM, Forsythe IJ, Gorvad AE;
 XX Griffin JA, Hafalia AJA, Ho A, Ison CH, Kable AE, Khare R, Lal PG;
 XX Lee S, Lee EA, Lee SY, Lehr-Wason PM, Li JX, Lindquist EA, Luo W;
 XX Marquis JP, Ramkumar J, Richardson TW, Sprague WW, Swarnakar A;
 XX Tang YT, Warren BA, Yang J, Yue H, Zebardjian Y, Zheng W;
 XX
 XX WPI; 2003-403125/38.
 XX
 XX N-PSDB; ACC44316.
 XX
 XX New human structural and cytoskeleton-associated proteins (SCAP) useful
 XX for diagnosing, treating and preventing diseases or conditions associated
 XX with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.
 XX
 XX Claim 1; Page 261-264; 361pp; English.
 XX
 XX This sequence represents a novel isolated human structural and
 XX cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and
 XX polynucleotides encoding them are useful in diagnosing, treating and
 XX preventing diseases or conditions associated with the decreased
 XX expression or over expression of SCAP, such as cell proliferative (e.g.
 XX cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
 XX disease, stroke), heart (e.g. hypertension, heart failure, angina) and
 XX skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral
 XX infections. These are also useful in assessing the effects of exogenous
 XX compounds on the expression of nucleic acid and amino acid sequences of
 XX SCAP. The SCAP or its fragments are useful in screening compounds for
 XX effectiveness as agonist or antagonist of the polypeptides, or in
 XX altering the expression of the target polynucleotide and compounds that
 XX specifically bind to or modulate the activity of the polypeptide. The
 XX microarray is useful in monitoring or measuring protein-protein
 XX interactions, drug-target interactions, and gene expression profiles
 XX
 XX Sequence 1258 AA;
 XX
 XX Query Match 12.8%; Score 134.5; DB 6; Length 1258;
 XX Best Local Similarity 23.4%; Pred. No. 0.0035;
 XX Matches 50; Conservative 43; Mismatches 106; Indels 15; Gaps 5;
 XX
 XX 2 SKKGLSASEKTRMMEISETKDVPQLKLEKIPAKGKITAMSVKVLQSLVDGMDV 61
 XX 259 ARKEAEKAEKRYMEWADTADAEATLQKWAERASLQOEVEKALKERVDELITD 318
 XX
 XX 62 CERI-----GTSNYWAFPPSKALHAR-KHKLVLESLSGSGKSHASLQKSIEKA 110
 XX 319 LEITKAEIEBKSGDGAASSYQLKQLEQNALKDALVRMRDLSSEKQEHVKLQKMEK- 377
 XX 111 KIGRCET--BERTLAKELSSLRDQEQKAEVKKYKCDPQVVEIRQANKVAEAAAR 168

Db 378 KNQLEVVVRQQRRLQBELSQAEISTIDELKEQVDAALGAE-ENVEMLTDRNLNLEKVKRE 436
 QY 169 WTDNIFAIAKSWAKKFGFEENKIDRTFGIPEDFD 202
 Db 437 LRETVGDLAEMNEMNDELQENARETELELREQLD 470
 RESULT 13
 ABR57415
 ID ABR57415 standard; protein; 1266 AA.
 XX
 XX ABR57415;
 XX
 XX 15-SEP-2003 (first entry)
 XX
 XX Human NOV2 protein SEQ ID NO:8.
 XX
 XX Human; NOVX; cytostatic; cardiatic; antiinflammatory; immunosuppressive;
 KW human; NOVX; cytostatic; cardiatic; antiinflammatory; immunosuppressive;
 KW antiarthritic; haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic;
 KW anorectic; antiasthmatic; nephrotropic; antiarthritic; hepatotropic;
 KW neuroprotective; nootropic; antibacterial; virucide; antiparasitic;
 KW relaxant; anticonvulsant; hypotensive; vasotropic; antiparkinsonian;
 KW vulnerary; angiogenic; antidiabetic; hypertension; diabetes; cancer;
 KW cardiomyopathy; atherosclerosis; hypertension; diabetes; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; obesity; asthma;
 KW acquired immunodeficiency syndrome; nephropathy; cirrhosis; arthritis;
 KW Alzheimer's disease; Parkinson's disease; Goitre; infection; stroke;
 KW muscular dystrophy; epilepsy; wasting disorder; chromosome 2.
 XX
 XX Homo sapiens.
 OS
 XX W0200294870-A2.
 XX
 XX 28-NOV-2002.
 XX
 XX 02-NOV-2001; 2001WO-US051580.
 XX
 XX 02-NOV-2000; 2000US-0245291P.
 XX
 XX 02-NOV-2000; 2000US-0245317P.
 XX
 XX 07-NOV-2000; 2000US-0246562P.
 XX
 XX 08-NOV-2000; 2000US-0246871P.
 XX
 XX 26-JAN-2001; 2001US-0264389P.
 XX
 XX 26-JAN-2001; 2001US-0264423P.
 XX
 XX 29-JAN-2001; 2001US-0264799P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Grosse WM, Macdougall JR, Smithson G, Millet I, Stone DJ;
 XX Gunther E, Ellerman K, Alsobrook JP, Lepley DM, Burgess CB;
 XX Spytek KA, Edinger SR, Gangolli EA, Gorman L, Taupier RJ, Li L;
 XX Guo X, Fernandes ER, Vernet CAM, Tcherev VI, Casman SJ, Shenoy S;
 XX Mishra V, Furtak K, Baumgartner JC, Colman SD;
 XX
 XX WPI; 2003-140359/13.
 XX
 XX N-PSDB; ACF03550.
 XX
 XX New NOVX polypeptide useful for preventing or treating NOVX-associated
 XX disorders, e.g. cancer, cardiomyopathy, atherosclerosis or diabetes, and
 XX in chromosome mapping, tissue typing or pharmacogenomics.
 XX
 XX Claim 1; Page 24; 346pp; English.
 XX
 XX ACF03547 to ACF03570 encode the human NOVX proteins (I) given in ABR57412
 XX to ABR57435 (I) have cytostatic, cardiatic, antiinflammatory, nootropic,
 XX immunosuppressive, antiarthritic, haemostatic, anti-HIV, antidiabetic,
 XX antiarteriosclerotic, anorectic, antiasthmatic, nephrotropic, virucide,
 XX antiarthritic, hepatotropic, neuroprotective, antibacterial, relaxant,
 XX vulnerary, angiogenic and antidiabetic activities, and can be used in
 XX gene therapy and vaccines. The NOVX polypeptides and their antibodies can
 XX be used to determine the presence or absence of (I) in a sample. The NOVX
 XX polypeptides, polynucleotides encoding them, and antibodies against them,
 XX are useful in manufacturing a medicament for treating or preventing a


```

PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX N-PSDB; AA159601.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 5376; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA158642-AA162213) with neurotropic,
CC immunosuppressant and cytototoxic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 1302 AA;
XX
Query Match 12.8%; Score 134.5; DB 4; Length 1302;
Best Local Similarity 23.4%; Pred. No. 0.0037;
Matches 50; Conservative 43; Mismatches 106; Indels 15; Gaps 5;
QY 2 SKKGLSAEKRTRMEIFSETKDVQLKDLKIAPEKGTITAMSVKVLQSLVDGMVD 61
DB 300 ARKEAKEALBAKERYMEADTADAIEMATLDKXMAEERAEISLQOEVEALKERVDELTTD 359
QY 62 CER1-----GTSNYWAFPSKALHAR-KHKLEVLSEQLSEGSQKHSIQKSIEKA 110
DB 360 LEILKABIEKSGDGAASSVQLKLEQNAHLKDALVRDILSSSEKQSHVQLKLMK- 418
QY 111 KIGRCET--BERTLAKELSSLRDQEQKAEVYKDCDPQVVEIRQANKVAKFAANR 168
DB 419 KNQLEVVROQRERLQELSQAEISTIDELKEQVDAALGAE-EMVEMLTDLNINLSEKVR 477
QY 169 WTDNIFAISWAKRGKGFENKIDRTFGIPEDFD 202
DB 478 LRETVGDLAEMNDELQENARETELEURQLD 511
RESULT 16
ADE15628
ID ADE15628 standard; protein; 266 AA.
XX
AC ADE15628;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human structural and cytoskeleton-associated protein (SCAP) #7.
XX

```

```

KW human; structural and cytoskeleton-associated protein; SCAP;
KW arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; myelofibrosis;
KW psoriasis; cancer; pneumonia; chronic bronchitis; yellow fever;
KW influenza; measles; mumps; HIV; human T lymphotropic virus; rabies;
KW gastroenteritis; encephalitis; rubella; epilepsy;
KW ischaemic cerebrovascular disease; stroke; cerebral neoplasm;
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
KW Parkinson's disease; amyotrophic lateral sclerosis; atrophy;
KW hereditary ataxia; multiple sclerosis; meningitis; brain abscess;
KW prion disease; Creutzfeldt-Jakob disease; insomnia; neurofibromatosis;
KW cerebral palsy; myasthenia gravis; anxiety.
XX
XX Homo sapiens.
OS
XX WO2003062391-A2.
PN
XX 31-JUL-2003.
PD
XX
XX 16-JAN-2003; 2003WO-US001772.
PF
XX
XX 18-JAN-2002; 2002US-0350702P.
PR
XX 25-JAN-2002; 2002US-0351715P.
PR 15-FEB-2002; 2002US-0357402P.
PR 10-MAY-2002; 2002US-0379880P.
PR 17-MAY-2002; 2002US-0381599P.
PR 07-JUN-2002; 2002US-0387270P.
PR 19-JUL-2002; 2002US-0397125P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Griffin JA, Richardson TW, Tang YT, Thangavelu K;
PI Forsythe IJ, Becha SD, Chawla NK, Hafalia AJA, Swarnakar A;
PI Marquis JP, Gorvad AE, Baugman MR, Lu DAM, Arvizu CS, Kable AE;
PI Lee SY, Ramkumar J, Jiang X, Jackson AA, Khare R, Elliott VS;
PI Bulloch SA, Xu Y, Lee S, Lehr-Mason PM;
XX
XX WPI; 2003-671469/63.
DR N-PSDB; ADE15659.
XX
XX New isolated polypeptides useful for treating e.g. cell proliferative
PT disorders, viral infections and neurological disorders.
XX
XX Claim 1; SEQ ID NO 7; 357pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC structural and cytoskeleton-associated proteins (SCAP). The SCAP DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of: arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,
CC myelofibrosis, psoriasis, primary cancer, pneumonia, chronic bronchitis,
CC yellow fever, influenza, measles, mumps, HIV, human T lymphotropic virus,
CC rabies, gastroenteritis, encephalitis, rubella, epilepsy, ischaemic
CC cerebrovascular disease, stroke, cerebral neoplasm, Alzheimer's disease,
CC Pick's disease, Huntington's disease, dementia, Parkinson's disease,
CC amyotrophic lateral sclerosis, atrophy, hereditary ataxias, multiple
CC sclerosis, meningitis, brain abscess, prion disease, Creutzfeldt-Jakob
CC disease, insomnia, neurofibromatosis, cerebral palsy, myasthenia gravis,
CC anxiety. The present amino acid sequence represents a human SCAP of the
CC invention.
XX
XX Sequence 266 AA;
SQ
Query Match 12.0%; Score 125.5; DB 7; Length 266;
Best Local Similarity 23.0%; Pred. No. 0.003;
Matches 51; Conservative 41; Mismatches 69; Indels 61; Gaps 8;
QY 2 SKKGLS-AEKRTRMEIFSETKDV-----FOLKDLKIAPEK 39
DB 45 AKKLLRVSEDRVLEELHKAEDSLLAEEAAKAAADVASLNRRRIQLVSELDRAQE 104
QY 40 KGITAMSVKEVLQSLVDD---GMVDCERIGTSTNYWAFPSKALHAR-----KHKLEVLSEQ 92
DB 105 RLATALQKLEAEAKADESERGM-----KVIESAQKDEEKMEIQEIQ 147

```

QY 93 LSEGSQKASLQKSIKAKI--GRCTEETRIAKELSSLRDQREOLKAEVVKYKCDPQ 150
Db 148 LKKAHIESLEAEAEALSEKGC-----AELEELKTVTNLKSLEAQAEKYSQKEDR 202
QY 151 VBEIRQANKVAEAKNRWTDNIFAIKSWAKRFGFEENKID 192
Db 203 YESEIKVLSKLEAEATR-----AEFAERSVTKLEKSID 236

RESULT 17
ABP73809
ID ABP73809 standard; protein; 1881 AA.
XX AC ABP73809;
XX DT 30-JAN-2003 (first entry)
XX DE DE Candida albicans essential protein SEQ ID NO 7646.
XX KW Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal.
XX OS Candida albicans.
XX PN WO200253728-A2.
XX PD 11-JUL-2002.
XX PF 26-DEC-2001; 2001WO-US049486.
XX PR 29-DEC-2000; 2000US-0259128P.
XX PR 20-FEB-2001; 2001US-00792024.
XX PR 22-AUG-2001; 2001US-0314050P.
XX PR (ELIT-) ELITRA PHARM INC.
XX PA Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX PI WPI; 2002-566694/60.
XX DR N-PSDB; ABZ32359.
XX FT Constructing strains for identifying gene products as effective targets
FT for therapeutic intervention, by inactivating in the strain one allele of
FT a gene and placing other allele of the gene under conditional expression.
XX PS Claim 44; SEQ ID NO 7646; 167pp + Sequence Listing; English.
XX CC The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans protein used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office

XX SQ Sequence 1881 AA;
Query Match 11.6%; Score 121.5; DB 5; Length 1881;
Best Local Similarity 23.8%; Pred. No. 0.095;
Matches 49; Conservative 43; Mismatches 79; Indels 35; Gaps 8;
QY 3 KKKLSAEKTRMWEIPSETKQVDFQKLEKIAPEKKGIT--AMSVKEVLQSLVDDGMV 60
Db 1670 KTKNSDTESKLEKQLELEKVK-----SDLQADEKLKIGITERIALKSELETVKNSG-- 1722
QY 61 DCEIRGTSNYWAPPS--KALHARKHKLLEVI-----ESQLSEGSOKHASLOKSI----- 107
Db 1723 ----LSTTSELAALTQVKSLEKEKEKELOFLSGNKSKELEDIYIQHSDISEKALTDEL 1778
QY 108 -EKAKIGRCETEETRIAKELSSLRDQREOLKAEVVKYKCDPQVVEIRQANKVAKEAA 166
Db 1779 KEKTKQFDDSKKLTLENDLTSTKLETEKTKTSKFKNLEERKKEIVKLKELELLK 1838
QY 167 NRWTDNIFAIKSWAKRFGFEENKID 192
Db 1839 N---DN-----SGAKKELSEKVSKELE 1856

RESULT 18
ABB66937
ID ABB66937 standard; protein; 1798 AA.
XX AC ABB66937;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 27603.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL11040.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX PS Disclosure; SEQ ID NO 27603; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1798 AA;

Query Match 11.5%; Score 120.5; DB 4; Length 1798;
Best Local Similarity 22.0%; Pred. No. 0.11;
Matches 48; Conservative 55; Mismatches 70; Indels 45; Gaps 9;
QY 3 KKKGLSABE--KRTMMEI--FSETKDV-----FOLKLEKIAPEKIGTAMSVKVL 51
DB 413 QKSKSELEMTKLTNNKEVELELKKVLGKETLLYENKOFKIELKGL--TQELIGLL 471
QY 52 QSL---VDDGMVDCERTGTSNYWAFPSKAL-----HARKHKLKYLE-- 90
DB 472 QAREKEVHDLLEIQLTATTSEQYYSKEVKDLKTELENEKLNKTELTSCHNKLLENKELT 531
QY 91 -----SOLSEGSQKHASLOKSIKAKIGRCETETETRIKAKELSLRDQREQ 136
DB 532 QETSDMTIELNQOEDINNKKOEERMLKQIE--NLQETETQLNELEYVREELKQKRD 589
QY 137 LKARVEKYKD--CD-----POVVEIRQANKVAKE--AANRWTDNIFAIRK----- 177
DB 590 VKCLDKSENCNLRKQVENKNYIELOENKALKKKGTAKSKQLNVYKVKLELE 649
QY 178 -SWAKRKFG-----FEENKI 191
DB 650 LESAKQKFGKITDTYQKEIEDKKI 673

RESULT 19
AAG66581
ID AAG66581 standard; protein; 975 AA.
XX AC AAG66581;
XX DT 22-OCT-2001 (first entry)
XX DE Human SCP-1 mutein.
XX KW Human; SCP-1; synaptonemal complex protein 1; cytostatic; gene therapy;
XX KW Human; mutant; cell transformation marker; diagnosis; cancer.
XX OS Homo sapiens.
XX FN US6232460-B1.
XX PD 15-MAY-2001.
XX PF 25-JUN-1998; 98US-00104324.
XX PR 15-JUL-1997; 97US-00892702.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Tuereci O, Sahin U, Pfreundschuh M;
XX DR WPI; 2001-342776/36.
XX PT Novel nucleic acid molecule encoding a mutein of SCP-1, useful as a
XX PT marker for cell transformation, and for identifying substances which are
XX PT immunoreactive and indicative of pathological conditions.
XX PS Example 2; Col 15-20; 11pp; English.
XX CC The invention relates to an isolated nucleic acid molecule which encodes
XX CC a mutein of synaptonemal complex protein 1 (SCP-1). The mutein comprises
XX CC a sequence of 976 amino acids fully defined in the specification (with
XX CC the proviso that His at position 225 is replaced by Phe, and Gly at
XX CC position 226 is replaced by Gln). The SCP-1 mutein is useful as a marker
XX CC for cell transformation, for diagnosis and treatment of cancer and for
XX CC screening substances which are immunoreactive and indicative of
XX CC pathological conditions. The present sequence is the mutein of SCP-1
XX CC provided in the specification
XX SQ Sequence 795 AA;

Query Match 11.2%; Score 117; DB 5; Length 795;
Best Local Similarity 23.9%; Pred. No. 0.079;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;
QY 3 KKKGLSABE--KRTMMEI--FSETKDV-----FOLKLEKIAPEKIGTAMSVKVL 51
DB 413 QKSKSELEMTKLTNNKEVELELKKVLGKETLLYENKOFKIELKGL--TQELIGLL 471
QY 52 QSL---VDDGMVDCERTGTSNYWAFPSKAL-----HARKHKLKYLE-- 90
DB 472 QAREKEVHDLLEIQLTATTSEQYYSKEVKDLKTELENEKLNKTELTSCHNKLLENKELT 531
QY 91 -----SOLSEGSQKHASLOKSIKAKIGRCETETETRIKAKELSLRDQREQ 136
DB 532 QETSDMTIELNQOEDINNKKOEERMLKQIE--NLQETETQLNELEYVREELKQKRD 589
QY 137 LKARVEKYKD--CD-----POVVEIRQANKVAKE--AANRWTDNIFAIRK----- 177
DB 590 VKCLDKSENCNLRKQVENKNYIELOENKALKKKGTAKSKQLNVYKVKLELE 649
QY 178 -SWAKRKFG-----FEENKI 191
DB 650 LESAKQKFGKITDTYQKEIEDKKI 673

RESULT 20
AAG66581
ID AAG66581 standard; protein; 976 AA.
XX AC AAG66581;
XX DT 22-OCT-2001 (first entry)
XX DE Human SCP-1 mutein.
XX KW Human; SCP-1; synaptonemal complex protein 1; cytostatic; gene therapy;
XX KW Human; mutant; cell transformation marker; diagnosis; cancer.
XX OS Homo sapiens.
XX FN US6232460-B1.
XX PD 15-MAY-2001.
XX PF 25-JUN-1998; 98US-00104324.
XX PR 15-JUL-1997; 97US-00892702.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Tuereci O, Sahin U, Pfreundschuh M;
XX DR WPI; 2001-342776/36.
XX PT Novel nucleic acid molecule encoding a mutein of SCP-1, useful as a
XX PT marker for cell transformation, and for identifying substances which are
XX PT immunoreactive and indicative of pathological conditions.
XX PS Example 2; Col 15-20; 11pp; English.
XX CC The invention relates to an isolated nucleic acid molecule which encodes
XX CC a mutein of synaptonemal complex protein 1 (SCP-1). The mutein comprises
XX CC a sequence of 976 amino acids fully defined in the specification (with
XX CC the proviso that His at position 225 is replaced by Phe, and Gly at
XX CC position 226 is replaced by Gln). The SCP-1 mutein is useful as a marker
XX CC for cell transformation, for diagnosis and treatment of cancer and for
XX CC screening substances which are immunoreactive and indicative of
XX CC pathological conditions. The present sequence is the mutein of SCP-1
XX CC provided in the specification
XX SQ Sequence 976 AA;

Query Match 11.2%; Score 117; DB 5; Length 795;
Best Local Similarity 23.9%; Pred. No. 0.079;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;
QY 3 KKKGLSABE--KRTMMEI--FSETKDV-----FOLKLEKIAPEKIGTAMSVKVL 51
DB 413 QKSKSELEMTKLTNNKEVELELKKVLGKETLLYENKOFKIELKGL--TQELIGLL 471
QY 52 QSL---VDDGMVDCERTGTSNYWAFPSKAL-----HARKHKLKYLE-- 90
DB 472 QAREKEVHDLLEIQLTATTSEQYYSKEVKDLKTELENEKLNKTELTSCHNKLLENKELT 531
QY 91 -----SOLSEGSQKHASLOKSIKAKIGRCETETETRIKAKELSLRDQREQ 136
DB 532 QETSDMTIELNQOEDINNKKOEERMLKQIE--NLQETETQLNELEYVREELKQKRD 589
QY 137 LKARVEKYKD--CD-----POVVEIRQANKVAKE--AANRWTDNIFAIRK----- 177
DB 590 VKCLDKSENCNLRKQVENKNYIELOENKALKKKGTAKSKQLNVYKVKLELE 649
QY 178 -SWAKRKFG-----FEENKI 191
DB 650 LESAKQKFGKITDTYQKEIEDKKI 673

XX	Sequence 976 AA;
SQ	
Query Match	11.2%; Score 117; DB 6; Length 976;
Best Local Similarity	23.9%; Pred. No. 0.1;
Matches 63; Conservative	38; Mismatches 85; Indels 78; Gaps 13;
Qy	3 KKIGLSAE--KTRMMEI-FSETKV-----FQLKDLEKIAPKEKITAMSVKEVL 51
Db	413 QKSSELEENTKLNNKEVELEELKKVLGSEKTLTYENKQFEIAELKG-TEQELIGLL 471
Qy	52 QSL----VDDGWDCEIRIGTSNYWAPPKAL-----HARKHKLEVLE-- 90
Db	472 QAREKEVHDLEIQLTALTITSEQYYSKEVKDKLTELENEKLNKTELTSCHCNKLS-ENKELT 531
Qy	91 -----SQLSEGSOQHASLOKSIIEKAKIGRCETETERTRIASELSSIRDREQ 136
Db	532 QETSMDTLELKNOQEDINNKKQEERMLKOIE--NLQETETQLRNELEYVREELKQKRDE 589
Qy	137 LKAEEVKYD-CD-----PQVVEIRQANKVAKE--AANRWTDNIFAIK----- 177
Db	590 VKCKLDKSENCCNLRKOVENKNKYIELQQENKALKKKGTASKQLNVYEIKVNKLELE 649
Qy	178 -SWAKRKFG-----FEEKI 191
Db	650 LESAKQKFGIITDYQKEIEDKKI 673
RESULT 22	
ADC09595	ID ADC09595 standard; protein; 976 AA.
XX AC	ADC09595;
XX AC	
XX DT	18-DEC-2003 (first entry)
XX SC	SCP-1 #SEQ ID 596.
DE DE	
XX KW	Epitope; immunological; vaccine;
KW KW	major histocompatibility complex class I; MHC class I; cancer;
KW KW	immunisation.
XX OS	Unidentified.
XX XX	
PN PN	WO2003008537-A2.
XX PD	30-JAN-2003.
XX PF	29-MAR-2002; 2002WO-US010189.
XX PR	06-APR-2001; 2001US-0282211P.
PR PR	07-NOV-2001; 2001US-0337017P.
PR PR	07-MAR-2002; 2002US-0363210P.
XX PA	(CTLI-) CTL IMMUNOTHERAPIES CORP.
XX FI	Simard JUL, Diamond DC, Liu L, Xie Z;
XX DR	WPI; 2003-248010/24.
XX PT	Epitope having high affinity for major histocompatibility complex class I
PT PT	useful for treating an animal, evaluating immunogenicity of a vaccine or
PT PT	therapeutic composition and for diagnosing a disease.
XX PS	Claim 1; SEQ ID NO 596; 239pp; English.
XX CC	The invention relates to an isolated epitope polypeptide that has high
CC CC	affinity for major histocompatibility complex (MHC) class I, and an
CC CC	epitope cluster comprising the polypeptide, also disclosed is a vaccine
CC CC	or immunotherapeutic composition containing an epitope of the invention.
CC CC	Compositions of the invention may be used in the treatment of cancer. The
CC CC	method can be combined with a radiation therapy, chemotherapy,
CC CC	biochemotherapy or surgery. The composition is also useful for evaluating

CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.
XX
SQ Sequence 976 AA;
Query Match 11.2%; Score 117; DB 7; Length 976;
Best Local Similarity 23.9%; Pred. No. 0.1;
Matches 63; Conservative 39; Mismatches 85; Indels 78; Gaps 13;
QY 3 KKGLSABE--KTRMWEI--FSTKDV-----FQKDLKLEKIAPKKGITAMSKVLE 51
DB 413 QKSSSELEETKLTNNKEVELEELKKVLGKETLLYENKQFEKIABELKG--TEQELIGLL 471
QY 52 QSL---VDGWDCEIRIGTSNYWAPSKAL-----HAKHKLVELE-- 90
DB 472 QAREKEVHDLLEIQLTATTSEQYYSKEVKDLTELENEKLNKNTLTSHCNKLSLENKELT 531
QY 91 -----SOLSEGSQSHASLOKSIEKAKIGRCGETEERTRLAKELSSLRDQRQ 136
DB 532 QETSDMTLELNQOEDINNNKQOERMUKQIE--NLQETETQLRNELEVREELKQKQDE 589
QY 137 LKAEVEKYKD-CD-----PQVVEIRQANKVAKB--AARWTDNTPAK----- 177
DB 590 VKCKLDKSEENCNLRKQVENKKNKYIEELQOENKALKKKGTAEKQNLNYYEIKVKNKLELE 649
QY 178 -SWAKRFG-----PRENKI 191
DB 650 LSAKQKFGELTDTYQKEIEDKKI 673
RESULT 23
ABO14658
ID ABO14658 standard; protein; 1328 AA.
AC ABO14658;
XX
XX 25-AUG-2003 (first entry)
DE Novel human protein #31.
XX Human; NOV; gene therapy; endocrine related disease; diabetes;
KW metabolism-related disease; obesity; central nervous system disorder;
KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
KW stroke; infection.
XX
OS Homo sapiens.
XX
XX WO2003023002-A2.
XX
XX 20-MAR-2003.
XX
XX 09-SEP-2002; 2002WO-US028539.
XX
XX 07-SEP-2001; 2001US-0318120P.
XX 07-SEP-2001; 2001US-0318130P.
XX 10-SEP-2001; 2001US-0318430P.
XX 17-SEP-2001; 2001US-032636P.
XX 17-SEP-2001; 2001US-0322781P.
XX 17-SEP-2001; 2001US-0322816P.
XX 17-SEP-2001; 2001US-0322817P.
XX 19-SEP-2001; 2001US-0323519P.
XX 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 17-APR-2002; 2002US-0373212P.
PR 06-SEP-2002; 2002US-00236177.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Spytex KA, Patturajan M, Gorman L, Li L, Anderson DM, Zhong M;
PI Gerlach V, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
PI Shinkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
PI Rieger DK, Taupier RJ, Shency SG, Liu X, Padigaru M, Alsobrook JP;
PI Lepley DM, Edinger SR, Burgess CE;
XX
XX WPI; 2003-313242/30.
DR N-PSDB; ACD19351.
XX
PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
PT and polynucleotides, useful in gene therapy, e.g. for treating or
PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
PT stroke or infections.
XX
XX Claim 1; Page 151; 586pp; English.
XX
XX The invention describes a new isolated polypeptide (NOVX). The NOVX
XX polypeptide, nucleic acid and antibody are useful as therapeutics,
XX particularly in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, which includes a pathology associated
XX with NOVX polypeptide. The DNA encoding the protein is useful in gene
XX therapy for treating the disease or condition. In particular, the NOVX
XX polypeptide or polynucleotide is useful for treating endocrine/
XX metabolism-related diseases (e.g. obesity or diabetes), central nervous
XX system disorders (e.g. Alzheimer's disease, Parkinson's disease,
XX epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
XX and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
XX asthma, inflammatory bowel disease, rheumatoid arthritis or
XX osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
XX prostate or brain cancers, or melanoma), liver diseases (e.g. liver
XX cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
XX haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
XX These are also useful in developing powerful assay system for functional
XX analysis of various human disorders, as well as in diagnostic
XX applications, and for monitoring the effects of drugs during clinical
XX trials. This is the amino acid sequence of a novel human NOV protein
XX
SQ Sequence 1328 AA;
Query Match 11.1%; Score 116.5; DB 6; Length 1328;
Best Local Similarity 22.7%; Pred. No. 0.17;
Matches 49; Conservative 29; Mismatches 87; Indels 51; Gaps 5;
QY 3 KKGLSABEKRTRMEIFSTKDVFLQKLEKIAPKKGITA--MSVKEVLQSLVDDGMV 60
DB 1025 RKQNDLRKNWMEALASTKMLQDKVKNKTSKERQQQVEAVELEAKVLLKPKVSV 1084
QY 61 DCEIRIGTSNYWAPPSKALH-----ARKHKLVEISQLSEGSQSHASLOKSIEKAKIGRC 115
DB 1085 P-SNLISYGEWLHGFEKKAECMAGTSGSEVKVLEHKLKADENHLLQLECEKYSVLA 1143
QY 116 ETE-----ERTRLA-KELSSLRD 132
DB 1144 ETEGILQKLSRQVEENKKNVVDKSHKTIKQMQSFTSSQELRLSRSENKDIENLR 1203
QY 133 QREQLKAEVKYKCDPQVVEIRQANKVAKAANR 169
DB 1204 EREHLEMELEKAEEMRSTYTVTEVRELKQALNETLTK 1239
RESULT 24
ADE15650
ID ADE15650 standard; protein; 1329 AA.
XX

AD15650;
29-JAN-2004 (first entry)
Human structural and cytoskeleton-associated protein (SCAP) #29.
human; structural and cytoskeleton-associated protein; SCAP;
arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; myelofibrosis;
psoriasis; cancer; pneumonia; chronic bronchitis; yellow fever;
influenza; measles; mumps; HIV; human T lymphotropic virus; rabies;
gastroenteritis; encephalitis; rubella; epilepsy;
ischemic cerebrovascular disease; stroke; cerebral neoplasm;
Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
Parkinson's disease; amyotrophic lateral sclerosis; atrophy;
hereditary ataxia; multiple sclerosis; meningitis; brain abscess;
prion disease; Creutzfeldt-Jakob disease; insomnia; neurofibromatosis;
cerebral palsy; myasthenia gravis; anxiety.
Homo sapiens.
WO2003062391-A2.
31-JUL-2003.
16-JAN-2003; 2003WO-US001772.
18-JAN-2002; 2002US-0350702P.
25-JAN-2002; 2002US-035715P.
15-FEB-2002; 2002US-0357402P.
10-MAY-2002; 2002US-0379880P.
17-MAY-2002; 2002US-0381599P.
07-JUN-2002; 2002US-0387270P.
19-JUL-2002; 2002US-0397125P.
(INCY-) INCYTE GENOMICS INC.
Yue H, Griffin JA, Richardson TW, Tang YT, Thangavelu K;
Forsythe IU, Becha SD, Chawla NK, Hafalia AJA, Swarnakar A;
Marquis JP, Gorvad AE, Baughn MR, Lu DAM, Arvizu CS, Kable AE;
Lee SY, Ramkumar J, Jiang X, Jackson AA, Khare R, Elliott VS;
Bullock SA, Xu Y, Lee S, Lehr-Mason PM;
WPI: 2003-671468/53.
N-PSDB; AD15681.
New isolated polypeptides useful for treating e.g. cell proliferative
disorders, viral infections and neurological disorders.
Claim 1; SEQ ID NO 29; 357pp; English.
The invention comprises the amino acid and coding sequences of human
structural and cytoskeleton-associated proteins (SCAP). The SCAP DNA and
protein sequences of the invention are useful for the diagnosis and
treatment of: arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,
myelofibrosis, psoriasis, primary cancer, pneumonia, chronic bronchitis,
yellow fever, influenza, measles, mumps, HIV, human T lymphotropic virus,
rabies, gastroenteritis, encephalitis, rubella, epilepsy, ischaemic
cerebrovascular disease, stroke, cerebral neoplasm, Alzheimer's disease,
Pick's disease, Huntington's disease, dementia, Parkinson's disease,
amyotrophic lateral sclerosis, atrophy, hereditary ataxia, multiple
sclerosis, meningitis, brain abscess, prion disease, Creutzfeldt-Jakob
disease, insomnia, neurofibromatosis, cerebral palsy, myasthenia gravis,
anxiety. The present amino acid sequence represents a human SCAP of the
invention.
Sequence 1329 AA;
Query Match 11.1%; Score 116.5; DB 7; Length 1329;
Best Local Similarity 22.7%; Pred. No. 0.17;
Matches 49; Conservative 29; Mismatches 87; Indels 51; Gaps 60
3 KKKGLSAEEKTRMMEIFSETKDVQLKLEKIAPEKKGITA--MSVKEVLSQSLVDGHW 60

Best Local Similarity 19.7%; Pred. No. 0.38; Matches 46; Conservative 47; Mismatches 64; Indels 77; Gaps 8;

QY 1 MSKKKGLSABEKRTRMEIFSETKDVFLQKDLKLE--KIAPKEGITAMSVKVLQSLVDDG 58
 DB 1637 LQKERIQVLEDQRTQTKILEE-----DLQIKLSLRERGRELTTORQLMQRAEAG 1688
 QY 59 M-----VDCE-----RIGTSNYW 72
 DB 1689 KGPSKAQSGSLHMKLILRDKEKEVECCQHHIHELQELKDQLEQLGLHRKVGET---- 1744
 QY 73 AFPSKALHAKHKLKLVLESOLSE-----GSQKHASLQKSIKAKIGRCETETRLAKELS 128
 DB 1745 ---SLLSQREQEIVLVQQLQEARQGLKEQSLQSLQDEAQRALAQRDQ-----ELE 1795
 QY 129 SLRDQRLKAEVKYKDCDPQVVEIRQANKVAKAANRWTDNIPAIKSWAKR 182
 DB 1796 ALQEQEQQAQGBERVKEKADALQGALEQAHTLKERHGELODH---KEQARR 1845

RESULT 26
 AAY77575
 ID AAY77575 standard; protein; 2442 AA.
 AC AAY77575;
 DT 08-MAY-2000 (first entry)
 DE Human cytoskeletal protein (HCYT) (clone 1655208).
 KW Human cytoskeletal protein; HCYT; cell proliferation; immunological;
 KW reproductive; nervous disorder; cytostatic; immunosuppressive; anti-HIV;
 KW anti-diabetic; antiarteriosclerotic; dermatological; anti-inflammatory;
 KW anti-infertility; vasotropic; cardiant.
 OS Homo sapiens.
 PN WO200006730-A2.
 PD 10-FEB-2000.
 XX 30-JUL-1999; 99WO-US017167.
 XX 31-JUL-1998; 98US-0155185P.
 PR 04-AUG-1998; 98US-0160081P.
 PR 19-AUG-1998; 98US-0155228P.
 XX (INCY-) INCYTE PHARM INC.
 XX Bandman O, Tang YT, Yue H, Corley NC, Guegler KJ, Azimzai Y;
 PI Patterson C, Lal P, Baughn MR;
 DR WPI; 2000-195297/17.
 DR N-PSDB; AAZ58980.
 XX Human cytoskeletal proteins useful for diagnosing, treating preventing
 PT cell proliferative, immunological, reproductive, developmental and
 PT nervous disorders.
 XX Claim 1; Page 71-76; 84pp; English.
 PS The invention provides human cytoskeletal proteins (HCYT) and nucleic
 CC acids encoding the proteins. The HCYT polypeptides can be expressed by
 CC standard recombinant methodology. The HCYT polypeptides, modulators and
 CC antibodies are useful for treating or preventing a disorder associated
 CC with decreased and increased expression or activity of HCYT in mammals.
 CC The polypeptides are also useful for diagnosing HCYT activity disorders
 CC such as cell proliferative, immunological, reproductive, developmental
 CC and nervous disorders. Sequences AAY77569-576 represent HCYT polypeptides
 XX Sequence 2442 AA;
 SQ Query Match 11.1%; Score 116.5; DB 3; Length 2442;

Best Local Similarity 19.7%; Pred. No. 0.39; Matches 46; Conservative 47; Mismatches 64; Indels 77; Gaps 8;

QY 1 MSKKKGLSABEKRTRMEIFSETKDVFLQKDLKLE--KIAPKEGITAMSVKVLQSLVDDG 58
 DB 1662 LQKERIQVLEDQRTQTKILEE-----DLQIKLSLRERGRELTTORQLMQRAEAG 1713
 QY 59 M-----VDCE-----RIGTSNYW 72
 DB 1714 KGPSKAQSGSLHMKLILRDKEKEVECCQHHIHELQELKDQLEQLGLHRKVGET---- 1769
 QY 73 AFPSKALHAKHKLKLVLESOLSE-----GSQKHASLQKSIKAKIGRCETETRLAKELS 128
 DB 1770 ---SLLSQREQEIVLVQQLQEARQGLKEQSLQSLQDEAQRALAQRDQ-----ELE 1820
 QY 129 SLRDQRLKAEVKYKDCDPQVVEIRQANKVAKAANRWTDNIPAIKSWAKR 182
 DB 1821 ALQEQEQQAQGBERVKEKADALQGALEQAHTLKERHGELODH---KEQARR 1870

RESULT 27
 AAB96332
 ID AAB96332 standard; protein; 880 AA.
 AC AAB96332;
 DT 29-OCT-2001 (first entry)
 DE Putative P. abyssi ATPase involved in DNA repair #2.
 KW Hyperthermophilic archaeon; hyperthermophilic protein.
 OS Pyrococcus abyssi.
 XX Pyrococcus abyssi.
 PN FR2792651-A1.
 XX 27-OCT-2000.
 XX 21-APR-1999; 99FR-00005034.
 XX 21-APR-1999; 99FR-00005034.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (IFRE-) IFREMER INST FR RECH EXPL MER.
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
 PI Querellou J, Weissenbach J, Saurin W, Heilig R;
 DR WPI; 2001-126236/14.
 XX New nucleotide sequences isolated from Pyrococcus abyssi encode proteins
 PT useful in industry.
 PS Claim 7; Page 1003-1006; 1657pp; French.
 CC The present invention relates to the genomic sequence of Pyrococcus
 CC abyssi (see AAF6431 and AAH4123-7) and P. abyssi proteins. P. abyssi is
 CC a hyperthermophilic archaeon, which is isolated from deep-sea
 CC hydrothermal vents. The present sequence is one such P. abyssi protein.
 CC The proteins of the present invention have various potential industrial
 CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade. Note: This patent is in the same patent family as
 CC WO2000065062, which contains additional sequences as shown in AAB99132-
 CC AAB99143, AAH75903-AAH75920 and AAG66436
 XX Sequence 880 AA;
 SQ Query Match 11.0%; Score 115.5; DB 4; Length 880;
 Best Local Similarity 24.1%; Pred. No. 0.12;
 Matches 58; Conservative 38; Mismatches 78; Indels 67; Gaps 10;

Db 257 RKGLLEKIVQIRSEIEEKAKISELEIVQDIPKLOEKEVYRKLKGR----- 306

QY 54 LVDGMVDCERIGTSNYWAFPSKAL-----HARKHLEVLSELSQSGSKHASL----- 103

Db 307 --DEYESKLRLLEKELSKWESLKAIEEVYKEGKKKKAEEIREKLSIEKRLEELKPY 364

QY 104 -----QKSIEKAK-----IGRCET--ERTRLAKELSSL-----RD 132

Db 365 VEELEDAKQVQKQIERLKAHLKGLSPGEVIEKLESEKERTIEBAIKETTRIGQMOE 424

QY 133 QRELKA--EVEKYKCDPOVVEEIRQANKVAKAANRWTDNIFAIKSWAKKFGFPEENK 190

Db 425 KNEEMKAIEELRKAEGKCPVCGRELTEHK--KELMERYTLEIKLIEBELKETT-ERK 481

QY 191 I 191

Db 482 L 482

RESULT 28

ADB67124

ID ADB67124 standard; protein; 284 AA.

XX

AC ADB67124;

XX

DT 04-DEC-2003 (first entry)

XX

DE Tropomyosin beta chain cardiac muscle SEQ ID NO:144.

XX

KW staged assembly; nanostructure; peptide nucleic acid; PNA;

KW structural reinforcement; aerogel; paper; plastic; cement;

KW tensile strength; identification marker; anti-counterfeiting marker;

KW enzyme support; catalyst support; assembly scaffold; nanowire;

KW nanocircuit; molecular sieve; molecular filter; biosensor.

XX

OS Sus scrofa.

XX

PN WO2003072829-A1.

XX

PD 04-SEP-2003.

XX

PF 21-FEB-2003; 2003WO-US0005390.

XX

PR 21-FEB-2002; 2002US-00080608.

XX

PA (NANO-) NANOFAMES INC.

XX

PI Hyman PL, Goldberg EB;

XX

XX WPI; 2003-721788/68.

XX

PT Staged assembly of nanostructures, useful e.g. in biosensors or as

PT catalyst supports, using assembly units derived from peptide nucleic

PT acids.

XX

PS Disclosure; Page 55; 118pp; English.

XX

CC The present invention describes a method (M1) for the staged assembly of

CC a nanostructure using peptide nucleic acids (PNAs). M1 comprises: (a)

CC contacting a nanostructure intermediate (NSI) having at least one unbound

CC joining element (JE) with an assembly unit (AU) that comprises several

CC different JE where: (i) none of these JE can interact with itself or

CC other JE; and (ii) only one JE in AU and a single unbound JE in NSI are

CC complementary, so that AU becomes non-covalently linked to NSI to produce

CC a new NSI for use in subsequent cycles; (b) removing unbound AU; and (c)

CC cyclic repetition of (a) and (b) to form a nanostructure. The new feature

CC is that the complementary JE in at least one cycle are PNAs. Also

CC described are nanostructures formed from many AU, comprising different

CC JE, where at least one AU includes PNA. M1 is useful for producing

CC nanostructures with a very wide range of potential applications, e.g.

CC structural reinforcements (for aerogels, paper, plastics or cement,

CC particularly as long fibres to improve tensile strength); identification

CC (anti-counterfeiting) markers; enzyme or catalyst supports; assembly

CC scaffolds; for construction of nanowires or nanocircuits; size markers

CC for electron microscopy; molecular sieves and filters; substrates for

CC optical and other surface coatings; scaffolds for solubilising enzymes or

CC for trapping, protecting and delivering specific molecules; in high-

CC density computer memories; as artificial zeolite for absorbing ions from

CC water and for construction of new materials, including use in biosensors.

CC PNAs are more homogeneous than inorganic nanoparticles generally used to

CC form nanostructures, so will produce structures with predictable geometry

CC and stoichiometry. The present sequence represents a protein containing

CC coiled coil dimerisation sequences that can be used for structural

CC elements of assembly units, given in the exemplification of the present

XX invention.

XX Sequence 284 AA;

QY 1 MSKKKGLSAE---EKRTMMEIRSETKDVQPKDLEKTIAPKEGKITAMSVKEVLSLVDD 57

Db 73 LAEKATDAEDVASLNRRIOQLFEELDRAQ-----ERLA-----TALQKLEAEKAADE 122

QY 58 ---GMVDCERIGTSNYWAFPSKALHAR----KHLEVLSELSQSGSKH----- 100

Db 123 SERGM-----KVIESRAQDKBEKMEIQIQLKEA--KHIAEDADRKYE 163

QY 101 -----ASLOKSTEKAKI--GRCTEERTFLAKELSLRDRQOLKAEVKEYKDCD 148

Db 164 EVARKLVIESDLERAEEAEELSEKGC-----AELEBELKTVTNLKSLEAQAESKYSQKE 218

QY 149 PQVVEEIRQANKVAKAANRWTDNIFAIKSWAKKFGFPEENKID 192

Db 219 DKYEELIKVLSDKLKEAETR-----AEFAERSVTKLEKSID 254

RESULT 29

ADB67131

ID ADB67131 standard; protein; 284 AA.

XX

AC ADB67131;

XX

DT 04-DEC-2003 (first entry)

XX

DE Tropomyosin alpha chain smooth muscle SEQ ID NO:151.

XX

KW staged assembly; nanostructure; peptide nucleic acid; PNA;

KW structural reinforcement; aerogel; paper; plastic; cement;

KW tensile strength; identification marker; anti-counterfeiting marker;

KW enzyme support; catalyst support; assembly scaffold; nanowire;

KW nanocircuit; molecular sieve; molecular filter; biosensor.

XX

OS Rattus sp.

XX

PN WO2003072829-A1.

XX

PD 04-SEP-2003.

XX

PF 21-FEB-2003; 2003WO-US0005390.

XX

PR 21-FEB-2002; 2002US-00080608.

XX

PA (NANO-) NANOFAMES INC.

XX

PI Hyman PL, Goldberg EB;

XX

XX WPI; 2003-721788/68.

XX

PT Staged assembly of nanostructures, useful e.g. in biosensors or as

PT catalyst supports, using assembly units derived from peptide nucleic

PT acids.

XX

PS Disclosure; Page 57; 118pp; English.

XX The present invention describes a method (M1) for the staged assembly of
 CC a nanostructure using peptide nucleic acids (PNAs). M1 comprises: (a)
 CC contacting a nanostructure intermediate (NSI) having at least one unbound
 CC joining element (JE) with an assembly unit (AU) that comprises several
 CC different JE where: (i) none of these JE can interact with itself or
 CC other JE; and (ii) only one JE in AU and a single unbound JE in NSI are
 CC complementary, so that AU becomes non-covalently linked to NSI to produce
 CC a new NSI for use in subsequent cycles; (b) removing unbound AU; and (c)
 CC cyclic repetition of (a) and (b) to form a nanostructure. The new feature
 CC is that the complementary JE in at least one cycle are PNAs. Also
 CC described are nanostructures formed from many AU, comprising different
 CC JE, where at least one AU includes PNA. M1 is useful for producing
 CC nanostructures with a very wide range of potential applications, e.g.
 CC structural reinforcements (for aerogels, paper, plastics or cement,
 CC particularly as long fibres to improve tensile strength); identification
 CC (anti-counterfeiting) markers; enzyme or catalyst supports; assembly
 CC scaffolds; for construction of nanowires or nanocircuits; size markers
 CC for electron microscopy; molecular sieves and filters; substrates for
 CC optical and other surface coatings; scaffolds for solubilising enzymes or
 CC for trapping, protecting and delivering specific molecules; in high-
 CC density computer memories; as artificial zeolite for absorbing ions from
 CC water and for construction of new materials, including use in biosensors.
 CC PNAs are more homogeneous than inorganic nanoparticles generally used to
 CC form nanostructures, so will produce structures with predictable geometry
 CC and stoichiometry. The present sequence represents a protein containing
 CC coiled coil dimerisation sequences that can be used for structural
 CC elements of assembly units, given in the exemplification of the present
 CC invention.

XX Sequence 284 AA;

Query Match 10.8%; Score 113.5; DB 7; Length 284;
 Best Local Similarity 21.4%; Pred. No. 0.043;
 Matches 48; Conservative 48; Mismatches 78; Indels 53; Gaps 8;

QY 2 SKKGLSA-EKTRMWEIPSEKDFQKLEKIAPEKGTAMSVK-EVLQSLVDDGM 59
 DB 45 AKKLLASEDEDRVLBEELHKAEDSLLAAD-ETAKAADVASLNRQLVBEELDRAQ 103
 QY 60 VDCERTGTSNYWAFPSKALHARKHKLVEVLSQSGSK----- 99
 DB 104 ---ERLATALQKLEAEKKADESERGMKVIESRAQKDEERMEITQLEKAKHIAEDADR 160
 QY 100 -----HASLQKSTEKAKI--GRCEETEERLAKELSSLRDQRELKAEVEKYK 145
 DB 161 KYEEVARKLVIESDLERAEERAEELSEGKC-----AELEELKTVTNLKSLEAQAEKYS 215
 QY 146 DCDPQVVEETROANKYAKAEANRWTONIFAISWAKRKFGFENKID 192
 DB 216 QKEDKYEIEIKVLSDKLKAEATR-----ABFAERSVTKLEKSID 254

RESULT 30
 AAM78985
 ID AAM78985 standard; protein; 484 AA.

AC AAM78985;

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1647.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WC0200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.
 XX 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK52118.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 PT Claim 20; Page 3984-3985; 6221pp, English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX Sequence 484 AA;

Query Match 10.8%; Score 113; DB 4; Length 484;
 Best Local Similarity 21.4%; Pred. No. 0.096;
 Matches 55; Conservative 46; Mismatches 74; Indels 82; Gaps 10;

QY 9 AEEKTRMWEIPSEKDFQKLEKIAPEKGTAMSVK-EVLQSL- 54
 DB 13 AEEKGTQAGEI-HOLKMDLVKERNVLOKTIENLQEQURDKQKMS--SLKERVKSQ 69
 QY 55 -----VDDGMVDCERI-----GTSNYWAFPSKALHARKH 85
 DB 70 ADTNTDTALTLEEALEKERTIERLKEQRDREREKQBEIDNY-----KKDLKLEK 124
 QY 86 LEVLSQSGSK-----QKHS-----LOKSTEKAKIGCEETEERELA 124
 DB 125 VSLQGLDSEKASLLDLKEHASSLASSGLKKDSRLKTLLEIALEQKBECLKVESQKKA 184
 QY 125 KELS-----SLRDQRELKAEVEKYKCDPQVVEETROANKYAKAEANRWTD--NIF 174
 DB 185 HEALEARASPEMSDRIOHLEREITRYKDESSAKQAQAEVDRLLEILKEVEKNDKKKIA 244
 QY 175 AISKWAKRKFGFENKI 191
 DB 245 ELESLSRQVKDQNKV 261

Search completed: September 27, 2004, 08:39:11
 Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 08:36:32 ; Search time 19 Seconds
(without alignments)
557.017 Million cell updates/sec

Title: US_10-087-190-3

Perfect score: 1047

Sequence: 1 MSKKGLSAEKRTRMEIF.....FEENKIDRTFGIPDFDYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : (Issued Patents AA:*)
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:.*
5: /cgn2_6/ptodata/2/iaa/PCITUS_COMB.pep:.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	593.5	56.7	127	4	US-09-621-976-4959
2	117	11.2	976	3	US-09-104-324B-4
3	114	10.9	284	4	US-09-914-259-55
4	113.5	10.8	284	4	US-09-914-259-62
5	112	10.7	284	4	US-09-914-259-43
6	110	10.5	284	4	US-09-914-259-46
7	110	10.5	284	4	US-09-914-259-49
8	109	10.4	284	4	US-09-914-259-51
9	107	10.2	281	4	US-09-914-259-63
10	107	10.2	284	4	US-09-167-206-10
11	107	10.2	284	4	US-09-914-259-40
12	107	10.2	284	4	US-09-914-259-48
13	107	10.2	284	4	US-09-914-259-50
14	106.5	10.2	284	4	US-09-914-259-41
15	106	10.1	534	4	US-09-103-664A-2
16	106	10.1	1786	3	US-08-973-462-8
17	105.5	10.1	224	2	US-08-272-255-16
18	105.5	10.1	224	5	PCF-US95-08585-16
19	105.5	10.1	245	4	US-10-164-593-34
20	105	10.0	245	4	US-09-914-259-65
21	105	10.0	251	4	US-09-914-259-64
22	105	10.0	284	4	US-09-914-259-60
23	102	9.7	281	4	US-09-914-259-45
24	102	9.7	284	4	US-09-914-259-47
25	102	9.7	284	4	US-09-914-259-57
26	102	9.7	372	1	US-07-813-584A-3
27	102	9.7	372	1	US-08-330-515-3

28	102	9.7	1031	4	US-09-914-259-24	Sequence 24, Appl
29	101.5	9.7	2682	4	US-09-595-684B-31	Sequence 31, Appl
30	100.5	9.6	284	4	US-09-914-259-59	Sequence 59, Appl
31	100	9.6	1104	3	US-08-923-992A-4	Sequence 4, Appl
32	100	9.6	1164	3	US-08-923-992A-10	Sequence 10, Appl
33	99.5	9.5	588	4	US-08-714-741-42	Sequence 42, Appl
34	99.5	9.5	864	4	US-08-714-741-40	Sequence 40, Appl
35	99.5	9.5	1312	2	US-08-592-126-148	Sequence 148, Appl
36	99.5	9.5	1312	2	US-08-587-080-51	Sequence 51, Appl
37	99.5	9.5	1312	4	US-09-168-595-148	Sequence 148, Appl
38	99	9.5	245	4	US-09-914-259-44	Sequence 44, Appl
39	99	9.5	1164	3	US-08-923-992A-2	Sequence 2, Appl
40	98	9.4	284	4	US-09-914-259-52	Sequence 52, Appl
41	98	9.4	1098	3	US-08-923-992A-8	Sequence 8, Appl
42	98	9.4	1128	3	US-08-923-992A-6	Sequence 6, Appl
43	98	9.4	1196	4	US-09-107-532A-3944	Sequence 3944, Ap
44	98	9.4	2482	1	US-08-328-254-6	Sequence 6, Appl
45	98	9.4	3248	1	US-08-353-700-1	Sequence 1, Appl
46	98	9.4	3248	5	PCF-US95-16216-1	Sequence 1, Appl
47	97.5	9.3	284	4	US-09-914-259-53	Sequence 53, Appl
48	97.5	9.3	284	4	US-09-914-259-54	Sequence 54, Appl
49	97.5	9.3	472	4	US-09-166-350-17	Sequence 17, Appl
50	96.5	9.2	343	3	US-08-937-271-17	Sequence 17, Appl
51	96	9.2	683	6	5210183-3	Patent No. 5210183
52	96	9.2	825	4	US-09-540-824-26	Sequence 26, Appl
53	95.5	9.1	431	4	US-09-286-981B-3	Sequence 3, Appl
54	95.5	9.1	967	4	US-09-914-259-21	Sequence 21, Appl
55	95	9.1	641	4	US-09-167-206-4	Sequence 4, Appl
56	94.5	9.0	871	4	US-09-134-001C-3979	Sequence 3979, Ap
57	94.5	9.0	3878	4	US-09-914-259-11	Sequence 11, Appl
58	94	9.0	174	4	US-09-198-452A-750	Sequence 750, Appl
59	94	9.0	935	4	US-09-914-259-25	Sequence 25, Appl
60	93.5	8.9	284	4	US-09-914-259-39	Sequence 39, Appl
61	93.5	8.9	344	6	5210183-2	Patent No. 5210183
62	93.5	8.9	1231	4	US-08-714-741-41	Sequence 41, Appl
63	93	8.9	928	4	US-09-134-000C-6590	Sequence 6590, Ap
64	93	8.9	1886	4	US-08-938-105-3	Sequence 3, Appl
65	93	8.9	1939	4	US-09-310-187A-1	Sequence 1, Appl
66	92.5	8.8	342	4	US-09-107-532A-5664	Sequence 5664, Ap
67	92.5	8.8	411	4	US-09-253-701-1	Sequence 1, Appl
68	92.5	8.8	443	2	US-08-795-475-6	Sequence 6, Appl
69	92.5	8.8	704	4	US-09-370-838-191	Sequence 191, Appl
70	92.5	8.8	963	4	US-09-914-259-22	Sequence 22, Appl
71	92.5	8.8	1087	4	US-09-914-259-12	Sequence 12, Appl
72	92	8.8	294	4	US-09-103-664A-5	Sequence 5, Appl
73	92	8.8	351	1	US-08-402-217A-2	Sequence 2, Appl
74	92	8.8	351	1	US-08-700-178-2	Sequence 2, Appl
75	92	8.8	351	3	US-08-995-654-2	Sequence 2, Appl
76	92	8.8	1581	4	US-09-866-108A-15754	Sequence 15754, A
77	92	8.8	1695	4	US-09-866-108A-15753	Sequence 15753, A
78	91.5	8.7	608	2	US-08-736-770-1	Sequence 1, Appl
79	91	8.7	963	4	US-09-914-259-20	Sequence 20, Appl
80	91	8.7	2261	4	US-09-526-193A-1	Sequence 1, Appl
81	91	8.7	2954	4	US-09-150-867-1	Sequence 1, Appl
82	90.5	8.6	284	4	US-10-164-595-32	Sequence 32, Appl
83	90.5	8.6	323	4	US-09-710-693-2	Sequence 2, Appl
84	90.5	8.6	415	3	US-08-938-830-1	Sequence 1, Appl
85	90.5	8.6	415	3	US-09-020-222-1	Sequence 1, Appl
86	90.5	8.6	977	4	US-09-010-147B-18	Sequence 18, Appl
87	90.5	8.6	2101	1	US-08-466-390-4	Sequence 4, Appl
88	90.5	8.6	2101	1	US-08-470-950-4	Sequence 4, Appl
89	90.5	8.6	2101	1	US-08-467-781-4	Sequence 4, Appl
90	90.5	8.6	2101	1	US-08-195-487-4	Sequence 4, Appl
91	90.5	8.6	2101	2	US-08-483-924-4	Sequence 4, Appl
92	90.5	8.6	2101	3	US-09-452-294-1	Sequence 1, Appl
93	90.5	8.6	2101	5	PCF-US93-06160-4	Sequence 4, Appl
94	90.5	8.6	2154	2	US-08-841-349-4	Sequence 4, Appl
95	90.5	8.6	2154	4	US-09-431-184A-4	Sequence 4, Appl
96	90	8.6	284	4	US-10-164-595-70	Sequence 70, Appl
97	90	8.6	305	3	US-08-937-271-10	Sequence 10, Appl
98	90	8.6	700	4	US-09-107-532A-5094	Sequence 5094, Ap
99	89.5	8.5	284	4	US-08-914-479A-6	Sequence 6, Appl
100	89.5	8.5	606	4	US-08-477-831C-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-09-621-976-4959
; Sequence 4959, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4959
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4959

Query Match      56.7%; Score 593.5; DB 4; Length 127;
Best Local Similarity 89.6%; Pred. No. 1.2e-50;
Matches 120; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

QY 1 MSKKKGLSABEKEKTRMWEIFSETKOVFOLKLEKIAPEKKGITAMSVKVLQSLVDDGVV 60
DB 1 MSKKKGLSABEKEKTRMWEIFSETKOVFOLKLEKIAPEKKGITAMSVKVLQSLVDDGVV 60
QY 61 DCRIGTSNYWAFPPSKALHARKHKLVLQSLSEGSQKHSIQAQKIGRCETEER 120
DB 61 DCRIGTSNYWAFPPSKALHARKHKLVLQSLSEGSQKHSIQAQKIGRCETEER 117
QY 121 TRLAKELSLRDOR 134
DB 118 ----IKLSGMQEER 127

RESULT 2
US-09-104-324B-4
; Sequence 4, Application US/09104324B
; Patent No. 6232460
; GENERAL INFORMATION:
; APPLICANT: T reci, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
; FILE REFERENCE: And Methods For Identifying Pathogenic Markers In A Sample Of
; TITLE OF INVENTION: No. 6232460mal Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,324B
; FILING DATE: 25-June-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/892,702
; FILING DATE: 15-July-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6232460man D.
; REGISTRATION NUMBER: 30,946

```

```

; REFERENCE/DOCKET NUMBER: LUD 5491
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5959
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-104-324B-4

```

```

Query Match      11.2%; Score 117; DB 3; Length 976;
Best Local Similarity 23.9%; Pred. No. 0.0072;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKKGLSAEE--KTRMWEI--PSETKOV-----POLKLEKIAPEKKGITAMSVKVL 51
DB 413 QKKSSELEMTKLTNNKEVELEBKVLGKEFILLYENKQFKIAELKG--TEQELIGLL 471
QY 52 QSL---VDDGMVDCERIGTSTNYWAFPPSKAL-----HARKHKLVLVLE-- 90
DB 472 QAREKEVHDLQIQLTAITTTSEQYVSKEVKDLKTELENEKLTSHCNKLSLENKELT 531
QY 91 -----SQLSEGSQKHSIQAQKIGRCETEERTRLAKELSLRDQREQ 136
DB 532 QETSDMTLELKNQOEDINNKKOEERMLKQIE--N-QETETQLRNELEYVREELKQKRD 589
QY 137 LKAEREYKQ-CD-----PQVVEEIRQANKVAKE--AANRWTDNIPAIK----- 177
DB 590 VKCKLDKSENCNLRKQVENKNKIYEELQENKALKKKGTAESKQLNVYVEIKVKNKLELE 649
QY 178 -SWAKRFG-----FEENKI 191
DB 650 LESAKQKFGSEITDTYQKEIEDKKI 673

```

RESULT 3

```

US-09-914-259-55
; Sequence 55, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-914-259-55

```

```

Query Match      10.9%; Score 114; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0027;
Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;

QY 1 MSKKKGLSABE--EKRTRMWEIFSETKOVFOLKLEKIAPEKKGITAMSVKVLQSLVDD 57
DB 73 LAEKATDAEADVASLNRRIQLFEEELDRAQ-----ERLA-----TALQKLEAEKAADE 122
QY 58 ---GMVDCERIGTSTNYWAFPPSKALHAR----KHKLVLQSLSEGSQKH----- 100
DB 123 SERGM-----KVIESRAQKDEKKEIQEIQLKEA--KHIAEDADRYE 163
QY 101 -----ASLQKSIEKAKI--GRCETEERTRLAKELSLRDQREQOLKAEVKKYKDCD 148
DB 164 EVARKLVIIISDLERAERAEALSEGKC-----AELEELKTVTNNLKSLEAQAKEYSQKE 218
QY 149 PQVVEEIRQANKVAKEAANRWTDNIPAIKSWAKRFGFEENKID 192

```

[illegible]

```
Query Match      10.5%; Score 110; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0067;
Matches 48; Conservative 39; Mismatches 80; Indels 40; Gaps 8;

QY 1 MSKKKGLSAB---EKTRMMEIFSETKDVQFQ-----LKDLKIAPK-EKGITAMSV 47
DB 73 LAEKATDAEDVASLNRRIQVVEELDRAQLATALQKLEAEAEKAADESERGMKVIEN 132
QY 48 KEVLQSLVDGMVDCERIGTSNYWAPPSKALHARKHKLVLSELSQSGSKHASIQKSI 107
DB 133 R-----ALKDEKIELQEI-----QLKEAKHIAEADRKYEAEVARKLVIIEGDL 176
QY 108 EKAKIGRCETEER--TRLAKELSLDRQRLKAEVEKYKDCDPQVVEIRQANKVAKEA 165
DB 177 ERAE-ERAELESKCAELBELKTVTNNLKSLEAQAEKYSQKEDKYEEIKVLTDLKLEA 235
QY 166 ANRWTDNIFAISWAKRKFGFEENKID 192
DB 236 ETR-----AFAERTVAKLEKSID 254

RESULT 8
US-09-914-259-51
; Sequence 51, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 284
; TYPE: PRF
; ORGANISM: Xenopus laevis
US-09-914-259-51

Query Match      10.4%; Score 109; DB 4; Length 284;
Best Local Similarity 23.7%; Pred. No. 0.0083;
Matches 49; Conservative 37; Mismatches 81; Indels 40; Gaps 8;

QY 1 MSKKKGLSAB---EKTRMMEIFSETKDVQFQ-----LKDLKIAPK-EKGITAMSV 47
DB 73 LSDKATDAEDVASLNRRIQVVEELDRAQLATALQKLEAEAEKAADESERGMKVIEN 132
QY 48 KEVLQSLVDGMVDCERIGTSNYWAPPSKALHARKHKLVLSELSQSGSKHASIQKSI 107
DB 133 R-----ALKDEKIELQEI-----QLKEAKHIAEADRKYEAEVARKLVIIEGDL 176
QY 108 EKAKIGRCETEER--TRLAKELSLDRQRLKAEVEKYKDCDPQVVEIRQANKVAKEA 165
DB 177 ERAE-ERAELESKCAELBELKTVTNNLKSLEAQAEKYSQKEDKYEEIKVLTDLKLEA 235
QY 166 ANRWTDNIFAISWAKRKFGFEENKID 192
DB 236 ETR-----AFAERTVAKLEKSID 254

RESULT 9
US-09-914-259-63
; Sequence 63, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
```

```
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 281
; TYPE: PRF
; ORGANISM: Rattus norvegicus
US-09-914-259-63

Query Match      10.2%; Score 107; DB 4; Length 281;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSABEKTRMMEIFSETKDVQFQKDLKIAPEKIGITAMSVKEVLQSLVD--- 57
DB 73 LAEKATDAE-----ADVASLNRRI-QLVEEELDRAQLATALQKLEAEAEKAADESER 125
QY 58 GWDCERIGTSNYWAPPSKALHAR-----KHKLVLSELSQSGSKH----- 100
DB 126 GM-----KVIESRAQKDEKMEIQIQLKEA--KHIAEDADRKYEVA 166
QY 101 -----ASLOKSIEKAKI--GRCETEERTRLAKELSLDRQRLKAEVEKYKDCDPQV 151
DB 167 RKLVIIESDLERAERAEALSEGKC-----AELEBELKTVTNNLKSLEAQAEKYSQKEDKY 221
QY 152 VEEIRQANKVAKAANRWTDNIFAISWAKRKFGFEENKID 192
DB 222 EEEIKVLSDKLKEAETR-----AFAERSVTKLEKSID 254

RESULT 10
US-09-167-206-10
; Sequence 10, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, MeiJa
; TITLE OF INVENTION: Nlk1 PROTEIN AND Nlk1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nlk1 protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 284
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-167-206-10

Query Match      10.2%; Score 107; DB 4; Length 284;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSABEKTRMMEIFSETKDVQFQKDLKIAPEKIGITAMSVKEVLQSLVD--- 57
DB 73 LAEKATDAE-----ADVASLNRRI-QLVEEELDRAQLATALQKLEAEAEKAADESER 125
QY 58 GWDCERIGTSNYWAPPSKALHAR-----KHKLVLSELSQSGSKH----- 100
DB 126 GM-----KVIESRAQKDEKMEIQIQLKEA--KHIAEDADRKYEVA 166
QY 101 -----ASLOKSIEKAKI--GRCETEERTRLAKELSLDRQRLKAEVEKYKDCDPQV 151
DB 167 RKLVIIESDLERAERAEALSEGKC-----AELEBELKTVTNNLKSLEAQAEKYSQKEDRY 221
QY 152 VEEIRQANKVAKAANRWTDNIFAISWAKRKFGFEENKID 192
DB 222 EEEIKVLSDKLKEAETR-----AFAERSVTKLEKSID 254

RESULT 11
```

US-09-914-259-40
; Sequence 40, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-40

Query Match 10.2%; Score 107; DB 4; Length 284;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSAEKRTRMMEIFSETKQVFLKDKLEKIAPEKGITAMSVKEVLSQSLVDD--- 57
DB 73 LAEKKATDAE-----ADVASLNRR1-QLVEEELDRAQERLATALQKLEAEKAADESER 125
QY 58 GMDVDCERIGTSNYWAFPSKALHAR-----KHKLEVLSELSQSGSKH----- 100
DB 126 GN-----KVIESRAQKDEKMEIQIOLKEA--KHIAEDADRKYEEVA 166
QY 101 -----ASLQKSIEKAKI--GRCTEERTRLAKELSLRDQRLKAEVKEKYKDCDPQV 151
DB 167 RKLVIIESDLERAEERAEISGKGC-----AELEELKTVTNLKSLEAQAEKYSQKEDKY 221
QY 152 VEIRQANKVAKAANRWTDNIFAISWAKRKFGEENKID 192
DB 222 EEBIKVLSDKLKEATR-----AEFAERSVTKLEKSID 254

RESULT 12
US-09-914-259-48
; Sequence 48, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-914-259-48

Query Match 10.2%; Score 107; DB 4; Length 284;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSAEKRTRMMEIFSETKQVFLKDKLEKIAPEKGITAMSVKEVLSQSLVDD--- 57
DB 73 LAEKKATDAE-----ADVASLNRR1-QLVEEELDRAQERLATALQKLEAEKAADESER 125
QY 58 GMDVDCERIGTSNYWAFPSKALHAR-----KHKLEVLSELSQSGSKH----- 100
DB 126 GN-----KVIESRAQKDEKMEIQIOLKEA--KHIAEDADRKYEEVA 166
QY 101 -----ASLQKSIEKAKI--GRCTEERTRLAKELSLRDQRLKAEVKEKYKDCDPQV 151
DB 167 RKLVIIESDLERAEERAEISGKGC-----AELEELKTVTNLKSLEAQAEKYSQKEDKY 221
QY 152 VEIRQANKVAKAANRWTDNIFAISWAKRKFGEENKID 192
DB 222 EEBIKVLSDKLKEATR-----AEFAERSVTKLEKSID 254

US-09-914-259-41
; Sequence 41, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-914-259-41

Query Match 10.2%; Score 106.5; DB 4; Length 284;
Best Local Similarity 23.0%; Pred. No. 0.015;
Matches 49; Conservative 31; Mismatches 76; Indels 57; Gaps 7;

US-09-914-259-40
; Sequence 40, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-40

Query Match 10.2%; Score 107; DB 4; Length 284;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSAEKRTRMMEIFSETKQVFLKDKLEKIAPEKGITAMSVKEVLSQSLVDD--- 57
DB 73 LAEKKATDAE-----ADVASLNRR1-QLVEEELDRAQERLATALQKLEAEKAADESER 125
QY 58 GMDVDCERIGTSNYWAFPSKALHAR-----KHKLEVLSELSQSGSKH----- 100
DB 126 GN-----KVIESRAQKDEKMEIQIOLKEA--KHIAEDADRKYEEVA 166
QY 101 -----ASLQKSIEKAKI--GRCTEERTRLAKELSLRDQRLKAEVKEKYKDCDPQV 151
DB 167 RKLVIIESDLERAEERAEISGKGC-----AELEELKTVTNLKSLEAQAEKYSQKEDKY 221
QY 152 VEIRQANKVAKAANRWTDNIFAISWAKRKFGEENKID 192
DB 222 EEBIKVLSDKLKEATR-----AEFAERSVTKLEKSID 254

US-09-914-259-48
; Sequence 48, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-914-259-48

Query Match 10.2%; Score 107; DB 4; Length 284;
Best Local Similarity 22.6%; Pred. No. 0.013;
Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKKGLSAEKRTRMMEIFSETKQVFLKDKLEKIAPEKGITAMSVKEVLSQSLVDD--- 57
DB 73 LAEKKATDAE-----ADVASLNRR1-QLVEEELDRAQERLATALQKLEAEKAADESER 125
QY 58 GMDVDCERIGTSNYWAFPSKALHAR-----KHKLEVLSELSQSGSKH----- 100
DB 126 GN-----KVIESRAQKDEKMEIQIOLKEA--KHIAEDADRKYEEVA 166
QY 101 -----ASLQKSIEKAKI--GRCTEERTRLAKELSLRDQRLKAEVKEKYKDCDPQV 151

QY 7 LSABEKTRMMEIFSETKDVQFQKLEKIAPKEGKITAMSVKEVLSQVLDV---GMVDCE 63
DB 72 LSABENAKAESEVASLNRRIQQLVEELDRQALATQKLEBAEKAADSESGM--- 127
QY 64 RIGTSNYWAPSPKALHAR-----KHKLEVLSEQLSEGSQKH----- 100
DB 128 -----KVIENRAQDKBEKMEIQIOKLEA--KHIAEADRKYEERVARKLVIL 172
QY 101 -ASLQKSIKAKIGRCETEERTRLAKELSLRDQREOLKAEVEKYKDCDPQVVEIRQAN 159
DB 173 EGDLEAEAEAESELSKCAE--LEELKLVNTNEAKLEAEQAEKYKQEDKYEERIKVLT 229
QY 160 KVAEEANRWTDNIFAKSWAKRFGFEENKID 192
DB 230 DKLKEAETR-----AEFAERSVTKLKSID 254

RESULT 15

US-09-103-664A-2

; Sequence 2, Application US/09103664A

; Patent No. 6458358

; GENERAL INFORMATION:

; APPLICANT: University of Kentucky Research Foundation

; APPLICANT: Timonsey, John

; APPLICANT: Artushin, Sergey

; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus

; TITLE OF INVENTION: and Assays Therefor

; FILE REFERENCE: 50229-212

; CURRENT APPLICATION NUMBER: US/09/103,664A

; CURRENT FILING DATE: 1998-06-23

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 534

; TYPE: PRT

; ORGANISM: Streptococcus equi

US-09-103-664A-2

Query Match 10.1%; Score 106; DB 4; Length 534;

Best Local Similarity 20.5%; Pred. No. 0.038;

Matches 46; Conservative 37; Mismatches 77; Indels 64; Gaps 7;

QY 8 SABEKTRMMEIFSETKDVQFQK-----DLEKIAPKEGKITAMSVKEVLSQV 55
DB 253 SABASRDKAFVSKDLADKLAATAEAKLMENVGSLDRLVESAKGEMAKLAIEDQLTA 312
QY 56 DCGMVDCEIGTSNYWAPSPKALHARKHKLVEVLSQVSEGSQKHASLQK----- 105
DB 313 DKAKADAE-LAAANDTIALQTELEKAKTELAVSERLIESGKRETAELQKQKQDASKALV 371
QY 106 -----SIEK-----AKIGRCETEE-----RTRLAK-----ELSLRDQREQLKA 139
DB 372 ESQANVAELEKQKASDAKVAELEKEVEAKAEVADLKAQLAKKBEELAVKKEALEA 431
QY 140 EVEKYKDCDPQVVEIR-----QANKVAEEANR 168
DB 432 KIEELKKAHAEBSLKLEMLEKMKDHANADLOAEINRLKQELADR 475

RESULT 16

US-08-973-462-8

; Sequence 8, Application US/08973462B

; Patent No. 6191270

; GENERAL INFORMATION:

; APPLICANT: DRUILHE, PIERRE

; APPLICANT: DAUBESIES, PIERRE

; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES

; FILE REFERENCE: 0660-0125-0 PCT

; CURRENT APPLICATION NUMBER: US/08/973,462B

; CURRENT FILING DATE: 1998-02-06

; EARLIER APPLICATION NUMBER: PCT/FR96/00894

; EARLIER FILING DATE: 1996-06-12

; EARLIER APPLICATION NUMBER: FR 95/07007

; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match 10.1%; Score 106; DB 3; Length 1786;

Best Local Similarity 18.7%; Pred. No. 0.19;

Matches 47; Conservative 55; Mismatches 77; Indels 72; Gaps 9;

QY 2 SKKGLSAEERKTRMMEIFSETKD-----VFQKDLKLEKIAPKEGKITAMSVKEV 52
DB 1186 SKEETESIKDKKDVSLVVEEVQNDMDSEVKEVLELKNMEELMKD-----AVEINDITS 1241
QY 53 SLVDD-----GMVDCERIGTSNYWAPSPKALHARKHKLVEVLSQVSEGSQ-----KHASL 103
DB 1242 KLIETQELNEVEADLI-----KDMEXLKELEKALSDESKIEIIDAKDQTL 1286
QY 104 OKSIEKA-----KIGRCETEERTRLAKELSLRDQREOLKAEVEKYKDCDPQVVEI 155
DB 1287 EKVIEEHHDITTLDEVVELKDVEEDKIEK-VSDLKDLKLEEDILKEVKEIKELSEILEY 1345
QY 156 RQANKV-----AKEAANRWTDNIFAKSWAKRFGFEENKIDR 193
DB 1346 KELNTIETDILEEKKEIKDKHFKFEFEAEIKLEADILKEVSSLEVE-----EKKLEE 1401
QY 194 TFGIPEDFDYI 204
DB 1402 VHELKEVEHI 1412

RESULT 17

US-08-272-255-16

; Sequence 16, Application US/08272255

; Patent No. 5824859

; GENERAL INFORMATION:

; APPLICANT: Cashmore, Anthony R.

; APPLICANT: Ahmad, Margaret

; APPLICANT: Lin, Chentao

; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of

; TITLE OF INVENTION: Using the Same

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris

; STREET: One Liberty Place, 46th floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/272,255

; FILING DATE: 08-JUL-1994

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Leary Ph.D., Kathryn

; REGISTRATION NUMBER: 36,317

; REFERENCE/DOCKET NUMBER: UPN-1795

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 224 amino acids


```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-16

Query Match
10.1%; Score 105.5; DB 2; Length 224;
Best Local Similarity 22.3%; Pred. No. 0.013;
Matches 48; Conservative 36; Mismatches 70; Indels 61; Gaps 8;

QY 7 LSAREKTRMMEIFSETKDVFLKDLKLEKAPKEKIGITAMSVKEVLQSLVDD---GMVDC 63
Db 12 LAADTAARAAEADVASINRRIQLVEEELDRAQLERLATALQKLEAEAKADESERGM--- 67
QY 64 RIGTSNYWAFPSKALHAR---KHKLVLSEQLSEGSQKH----- 100
Db 68 -----KVIESRAQKDEEKWEIOEIQLKEA--KHIAEDADRKYEEVARKLVII 112
QY 101 -ASLQKSIIEKAKI--GRCTEETRLAKELSLDRQREQLKAEVKKDCDPQVVEIRQ 157
Db 113 ESDLERAEERAEELSEGKC-----AELEELKVTNNLSLEAQAEKYSQKEDKYEEIKV 167
QY 158 ANKVAKEAANRWTDNIFAISWAKRKFGPEENKID 192
Db 168 LSDKLKEATR-----AEFAERSVTKLEKSID 194

RESULT 18
PCT-US95-08565-16
; Sequence 16, Application PC/TUS9508565
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08565
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,255
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-08565-16
```

Query Match 10.1%; Score 105.5; DB 5; Length 224;

```
Best Local Similarity 22.3%; Pred. No. 0.013;
Matches 48; Conservative 36; Mismatches 70; Indels 61; Gaps 8;

QY 7 LSAREKTRMMEIFSETKDVFLKDLKLEKAPKEKIGITAMSVKEVLQSLVDD---GMVDC 63
Db 12 LAADTAARAAEADVASINRRIQLVEEELDRAQLERLATALQKLEAEAKADESERGM--- 67
QY 64 RIGTSNYWAFPSKALHAR---KHKLVLSEQLSEGSQKH----- 100
Db 68 -----KVIESRAQKDEEKWEIOEIQLKEA--KHIAEDADRKYEEVARKLVII 112
QY 101 -ASLQKSIIEKAKI--GRCTEETRLAKELSLDRQREQLKAEVKKDCDPQVVEIRQ 157
Db 113 ESDLERAEERAEELSEGKC-----AELEELKVTNNLSLEAQAEKYSQKEDKYEEIKV 167
QY 158 ANKVAKEAANRWTDNIFAISWAKRKFGPEENKID 192
Db 168 LSDKLKEATR-----AEFAERSVTKLEKSID 194

RESULT 19
US-10-164-595-34
; Sequence 34, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-34

Query Match 10.1%; Score 105.5; DB 4; Length 245;
Best Local Similarity 23.7%; Pred. No. 0.015;
Matches 54; Conservative 35; Mismatches 64; Indels 75; Gaps 11;

QY 8 SAEEKR-TRMMEIFSETK-----DV-----FQKDLKLEKAPKEKIGITAMSVKEVLQ 53
Db 23 AAEERAGTQLQRELDHERKLRETAADVASLNRRQLVAVEEELDRAQLERLATALQKLEAEK 82
QY 54 LVDD---GMVDCERIGTSNYWAFPSKALHAR-----KHKLVLSEQLSEGSQKH----- 100
Db 83 AADESERGM-----KVIESRAQKDEEKWEIOEIQLKEA--KHIAEDAD 123
QY 101 -----ASLQKSIIEKAKI--GRCTEETRLAKELSLDRQREQLKAEVKK 144
Db 124 RYEEVARKLVIIESDLERAEERAEELSEGKC-----AELEELKVTNNLSLEAQAEKY 178
QY 145 KCDPQVVEIRQANKVAKEAANRWTDNIFAISWAKRKFGPEENKID 192
Db 179 SQEDRYEEIKVLSDKLKEATR-----AEFAERSVTKLEKSID 218

RESULT 20
US-09-914-259-65
; Sequence 65, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Yakowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 65
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-914-259-65

Query Match
Best Local Similarity 10.0%; Score 105; DB 4; Length 245;
Matches 51; Conservative 38; Mismatches 69; Indels 64; Gaps 10;

QY 2 SKKXGLSABEK--RTRMEIFSETKDVQFQKLEKIAPEKIGITAMSVKVLQSLVDD-- 57
Db 30 SLORELDQERKURETAAADVASLNRI--QLVVEELDRAQERLATALQKLEAEKAADESE 88
QY 58 -GMVDCERIGTSNYWAPPSKALHAR-----KHKLVLESQSLSEGSKH----- 100
Db 89 RGM-----KVIESRAQKDEEKWEIQEIQKLEA--KHIAEADRKYEYEV 129
QY 101 -----ASLQKSIIEKAKI--GRCTEETRLAKELSLRDQERQKAEVEKYKDCDPQ 150
Db 130 ARKLVIIESDLERAERAEKSEGK-----AELEELKTVTNLKSLEAQAEKYSQKEDK 184
QY 151 VVEIRQANKVAKAANRWTDNIFAISWAKRKFGFEENKID 192
Db 185 YEIEIKVLSDKLEAETR-----ABFAERSVTKLEKSID 218

RESULT 21
US-09-914-259-64
; Sequence 64, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-914-259-64

Query Match
Best Local Similarity 10.0%; Score 105; DB 4; Length 251;
Matches 51; Conservative 38; Mismatches 69; Indels 64; Gaps 10;

QY 2 SKKXGLSABEK--RTRMEIFSETKDVQFQKLEKIAPEKIGITAMSVKVLQSLVDD-- 57
Db 30 SLORELDQERKURETAAADVASLNRI--QLVVEELDRAQERLATALQKLEAEKAADESE 88
QY 58 -GMVDCERIGTSNYWAPPSKALHAR-----KHKLVLESQSLSEGSKH----- 100
Db 89 RGM-----KVIESRAQKDEEKWEIQEIQKLEA--KHIAEADRKYEYEV 129
QY 101 -----ASLQKSIIEKAKI--GRCTEETRLAKELSLRDQERQKAEVEKYKDCDPQ 150
Db 130 ARKLVIIESDLERAERAEKSEGK-----AELEELKTVTNLKSLEAQAEKYSQKEDK 184
QY 151 VVEIRQANKVAKAANRWTDNIFAISWAKRKFGFEENKID 192
Db 185 YEIEIKVLSDKLEAETR-----ABFAERSVTKLEKSID 218

RESULT 22
US-09-914-259-60
; Sequence 60, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-914-259-45

Query Match
Best Local Similarity 9.7%; Score 102; DB 4; Length 281;
Matches 50; Conservative 34; Mismatches 71; Indels 64; Gaps 9;

QY 1 MSKXKGLSABEKTRMWEIFSETKDVQFQKLEKIAPEKIGITAMSVKVLQSLVDD-- 57
Db 73 LADKATDAES-----EVALNRI--QLVVEELDRAQERLATALQKLEAEKAADESE 125
QY 58 GMVDCERIGTSNYWAPPSKALHAR-----KHKLVLESQSLSEGSKH----- 100
Db 126 GM-----KVIESRAQKDEEKWEIQEIQKLEA--KHIAEADRKYEYEV 166
QY 101 -----ASLQKSIIEKAKI--GRCTEETRLAKELSLRDQERQKAEVEKYKDCDPQ 153
Db 167 RKLVIIEGDLERAERAEKSEGKAB-----LEELKTVTNLKSLEAQAEKYSQKEDK 223
QY 154 EIRQANKVAKAANRWTDNIFAISWAKRKFGFEENKID 192
```

QY 58 GMVDCERIGTSNYWAPPKALHAR-----KHLEVLSEQLSEGSQKH----- 100
Db 126 GM-----KVLENRAQKDEKWEIQEIQKEA--KHIAEADRKVEEVA 166
QY 101 -----ASLQKIEKAKTGRCTEETRIKLESLRQRLKAEVEKYKDCDPQVVE 153
Db 167 RKLVIIEGDLERAERAELESKCAE---LEBELKTVTNLKSLEAQAEKYSQKEDKYEE 223
QY 154 EIRQANKVAKEAANRWTDNIFAIKSWAKGKFGFPEENKID 192
Db 224 EIKVLTDLKEAETR-----AEFAERSVTKLEKSID 254

RESULT 26

US-07-813-584A-3
; Sequence 3, Application US/07813584A
; Patent No. 5352588
; GENERAL INFORMATION:
; APPLICANT: Fischetti, Vincent A.
; APPLICANT: Bessen, Debra E.
; TITLE OF INVENTION: No. 5352588el Immunoglobulin A Binding Protein
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kittie Murray
; STREET: 98 Cutter Mill Road
; CITY: Great Neck
; STATE: NY
; COUNTRY: USA
; ZIP: 11021

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/813,584A
FILING DATE: 19911224
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Kittie
REGISTRATION NUMBER: 30,246
REFERENCE/DOCKET NUMBER: RU-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-482-1990
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-813-584A-3

Query Match 9.7%; Score 102; DB 1; Length 372;

Best Local Similarity 23.5%; Pred. No. 0.058; Indels 34; Gaps 6;
Matches 42; Conservative 36; Mismatches 67;

QY 7 LSABEKTRRMWEIFSE--TKDVFLQKD----LEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
Db 81 INAEENKKNKLEANKLENENYKLDGIDALEKEKEDLKTTLAKTKE-----NEIS 133
QY 61 DCERIGTSNYWA--FPSKALHAKHKLVELESQSEGSQKHASLOKSEI----- 108
Db 134 EASRKGLSRDLSEASRTAKKLEAKHOKLEAKKLTGEGNVSEASRKGLSNDLEASRAK 193
QY 109 ---KAKIGRCETEETRIKLESLSLRD-----OREQLKAEVEKYKDCDPQVVEBIRQA 158
Db 194 KELEAKYOKLETDHOLAEKQKLEADYQVSETSEKGLSRDLSEASREANKKVTSELTA 252

RESULT 27

US-08-330-515-3
; Sequence 3, Application US/08330515
; Patent No. 5556944

Db 224 EIKVLTDLKEAETR-----AEFAERSVTKLEKSID 254
RESULT 24
US-09-914-259-47
; Sequence 47, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Coturnix coturnix japonica
US-09-914-259-47

Query Match 9.7%; Score 102; DB 4; Length 284;
Best Local Similarity 22.8%; Pred. No. 0.04; Indels 64; Gaps 9;
Matches 50; Conservative 34; Mismatches 71;

QY 1 MSKKKGLSABEKTRRMWEIFSEKDVFLQKLEKIAPEKGITAMSVKEVLQSLVDD--- 57
Db 73 LADKKATDAES-----EVALNRRI-QLVEEELDRAQERLATALQKLEAEKAADESER 125
QY 58 GMVDCERIGTSNYWAPPKALHAR-----KHLEVLSEQLSEGSQKH----- 100
Db 126 GM-----KVLENRAQKDEKWEIQEIQKEA--KHIAEADRKVEEVA 166
QY 101 -----ASLQKIEKAKTGRCTEETRIKLESLRQRLKAEVEKYKDCDPQVVE 153
Db 167 RKLVIIEGDLERAERAELESKCAE---LEBELKTVTNLKSLEAQAEKYSQKEDKYEE 223
QY 154 EIRQANKVAKEAANRWTDNIFAIKSWAKGKFGFPEENKID 192
Db 224 EIKVLTDLKEAETR-----AEFAERSVTKLEKSID 254

RESULT 25

US-09-914-259-57
; Sequence 57, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Coturnix coturnix japonica
US-09-914-259-57

Query Match 9.7%; Score 102; DB 4; Length 284;
Best Local Similarity 22.8%; Pred. No. 0.04; Indels 64; Gaps 9;
Matches 50; Conservative 34; Mismatches 71;

QY 1 MSKKKGLSABEKTRRMWEIFSEKDVFLQKLEKIAPEKGITAMSVKEVLQSLVDD--- 57
Db 73 LADKKATDAES-----EVALNRRI-QLVEEELDRAQERLATALQKLEAEKAADESER 125

GENERAL INFORMATION:
APPLICANT: Fischetti, Vincent A.
APPLICANT: Bessen, Debra E.
TITLE OF INVENTION: No. 5556944e1 Immunoglobulin A Binding Protein
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/330,515
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,584
FILING DATE: 24-DEC-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: RU-100.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-330-515-3

Query Match 9.7%; Score 102; DB 1; Length 372;
Best Local Similarity 23.5%; Pred. No. 0.058;
Matches 42; Conservative 36; Mismatches 67; Indels 34; Gaps 6;
QY 7 LSAREKTRMELFSE--TKDVPQLKD---LEKIAPEKKGITAMSVKEVLQSLVDDGVV 60
DB 81 INAEENKKLEAINKELNENIYKQGDIDALEKEKEDLTKTAKTTKE-----NEIS 133
QY 61 DCRIGTSNYWA--PFSKALHARKHKLVEQLSESGQKHASLQKSE-----108
DB 134 EASRKGLSRDLEASRTAKKELEAKHQKLEAKNKKLTGEGNQVSEASRKGLSNDLEASRAK 193
QY 109 ---KAKIGCEETEERTRLAKEISSLD-----QREQLKAEVEKYKDCDPQVVEIROA 158
DB 194 KELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLSRDLEASREANKVTSILTQA 252

RESULT 28
US-09-914-259-24
Sequence 24, Application US/09914259
Patent No. 6495336
GENERAL INFORMATION:
APPLICANT: Makowski, Lee
APPLICANT: Hyman, Paul
APPLICANT: Williams, Mark
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
CURRENT APPLICATION NUMBER: US/09/010-999
FILING DATE: 2000-11-21
NUMBER OF SEQ ID NOS: 180
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 1031

TYPE: PRT
ORGANISM: Strongylocentrotus purpuratus
US-09-914-259-24
Query Match 9.7%; Score 102; DB 4; Length 1031;
Best Local Similarity 21.5%; Pred. No. 0.23;
Matches 48; Conservative 43; Mismatches 86; Indels 46; Gaps 8;
QY 2 SKKGLSAREKTR-----MMEIFSETKDVQLKDLKLEKI-----APKEKG 41
DB 631 AKMKLSNIRETEGKKHLEDLSMLNE--EIVKRAAEIRLTQDQDKREEDKMQS 688
QY 42 ITAM--SVKEVLQSLVDDGWDVCERIGTSNYWAPPSPKALHARKHKLVEQLSEGSQK 99
DB 689 ATEMQASMEQESHDAHQQLANLRT-----INEKHQMEELKDVNQRTLQ 738
QY 100 HASLOKSTIEKAKIGRCETEERTRIAKELSSLRDQREQLKAEVEKYKDC---DPQVVEIR 156
DB 739 HEKLDYELKI---EAEAKAKRLSELQQFDRREQAKDLKGLBETVAKELQTLHLNR 795
QY 157 Q-----ANKVAKAANRWTDNIFAISWAKRKFGFEENKIDR 193
DB 796 KLFVSDIQNRVKKALEGGDRDDSGSQAKQKISFLENNLEQ 838

RESULT 29
US-09-595-684B-31
Sequence 31, Application US/09595684B
Patent No. 6544766
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Ohashi, Cara
APPLICANT: Sakowicz, Roman
APPLICANT: Vaisberg, Eugeni
APPLICANT: Wood, Kenneth
APPLICANT: Yu, Ming
TITLE OF INVENTION: Human kinesins and methods of producing
TITLE OF INVENTION: and purifying human kinesins
FILE REFERENCE: cytop036
CURRENT APPLICATION NUMBER: US/09/595,684B
CURRENT FILING DATE: 2002-06-24
PRIOR FILING DATE: 09/295,612
PRIOR FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 2662
TYPE: PRT
ORGANISM: Human
US-09-595-684B-31

Query Match 9.7%; Score 101.5; DB 4; Length 2662;
Best Local Similarity 22.7%; Pred. No. 0.91;
Matches 46; Conservative 45; Mismatches 69; Indels 43; Gaps 9;
QY 1 MSKKKGLSAREK--TRMMEIFSETKDVQLKDLKLEKIAPEKKGITAMSVKEVLQSLVD-- 56
DB 1736 IDKLRGIVSEKTEINSNQKDLSEHNDALKAQDLK--IQEELRIAHHLKE-QQETIDKL 1792
QY 57 -----DGMVDCER-IGTSNYWAPPSPKALHARKHKLVEQLSEGSQKHAS----L 103
DB 1793 RGIVSEKTDKLSNNQKDLSEHNDALKAQDLK--IQEELRIAHHLKE-QQETIDKL 1851
QY 104 QKSIERAKIGRCETE-ERTELAKEL-----SSLRDQREQLKAEV 141
DB 1852 KQIKQKQSLTSLKLEINLNLQELHENLEEMKSVMKERDNLARVEETLKLERDQLKSSG 1911
QY 142 EKYKDCDPQVVEIROANKVAK 164
DB 1912 QETKARDLEIQOELKTARMLSK 1934

RESULT 30

```

US-09-914-259-59
; Sequence 59, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-914-259-59

Query Match          9.6%; Score 100.5; DB 4; Length 284;
Best Local Similarity 22.5%; Pred. No. 0.056; 77; Indels 57; Gaps 7;
Matches 48; Conservative 31; Mismatches 77;

QY      7  LSAEKKRTRMWEIFSETKDYFQKDLDEKIAPEKXGITAMSVKEVLQSLVDD---GMVDCE 63
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      72  LFAENNAKAESEVASLNRIQLVEEELDRAQERLATALQKLEAEAKADESERGM--- 127
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      64  RIGHTSNTYWAPPSKALHAR-----KHKLEVLSEQLSEGSQKH----- 100
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

DB      128 -----KVIENRAQKDEKMEIQIQLKEA--KHIAEADRKYYEEVARKLVII 172
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      101  -ASLQKSIEKAKIGRCETEERTRIAKLESSLRDQREOLKAEEVKYKCDPCQWVEEIRQAN 159
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

DB      173  EGDLERAEEAELSESACAE---LEEBLKVTNNLKSLQAQEKYSQKEDKYVEEIKVLT 229
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      160  KVAKEAANRWTDNTIFAIKSWAKRKFGFEENKID 192
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

DB      230  DKLKEATR-----AEPAERSVTKLEKSID 254
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: September 27, 2004, 08:38:05
Job time : 22 secs

[REDACTED]

SUMMARIES

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	1047	100.0	205	10	US-09-799-250-2	Sequence 2, Appli
2	1047	100.0	205	12	US-10-087-190-3	Sequence 3, Appli
3	1047	100.0	205	12	US-10-087-190-14	Sequence 14, Appl
4	1047	100.0	205	12	US-10-087-190-20	Sequence 20, Appl
5	1047	100.0	205	12	US-10-087-190-21	Sequence 21, Appl
6	1047	100.0	205	12	US-10-087-190-22	Sequence 22, Appl
7	1047	100.0	205	12	US-10-087-190-61	Sequence 61, Appl
8	1047	100.0	205	16	US-10-408-765A-1821	Sequence 1821, Ap
9	1036.5	99.0	206	12	US-10-087-190-44	Sequence 44, Appl
10	1036.5	99.0	206	12	US-10-087-190-45	Sequence 45, Appl
11	1011	96.6	198	12	US-10-087-190-24	Sequence 24, Appl
12	975	93.1	190	12	US-10-087-190-13	Sequence 13, Appl
13	975	93.1	190	12	US-10-087-190-19	Sequence 19, Appl
14	975	93.1	190	12	US-10-087-190-67	Sequence 67, Appl
15	975	93.1	190	12	US-10-087-190-68	Sequence 68, Appl

89 106.5 10.2 284 14 US-10-080-608A-41 Sequence 41, Appl
90 106.5 10.2 284 15 US-10-370-685-130 Sequence 130, App
91 106 10.1 243 15 US-10-104-047-3167 Sequence 3167, App
92 106 10.1 746 16 US-10-408-785A-1070 Sequence 1070, App
93 106 10.1 1786 9 US-09-742-096-3 Sequence 3, Appl
94 106 10.1 1787 12 US-10-415-253-2 Sequence 2, Appl
95 105.5 10.1 891 15 US-10-369-493-13477 Sequence 13477, A
96 105 10.0 245 14 US-10-080-608A-65 Sequence 65, Appl
97 105 10.0 245 15 US-10-370-685-154 Sequence 154, App
98 105 10.0 251 14 US-10-080-608A-64 Sequence 64, Appl
99 105 10.0 251 15 US-10-370-685-153 Sequence 153, App
100 105 10.0 284 14 US-10-080-608A-60 Sequence 60, Appl

ALIGNMENTS

RESULT 1
US-09-799-250-2
; Sequence 2, Application US/09799250
; Publication No. US2003032087A1
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121P1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799, 250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-2

Query Match 100.0%; Score 1047; DB 10; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60
QY 61 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSOKHASLOKSTIEKAKIGRCETEER 120
Db 61 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSOKHASLOKSTIEKAKIGRCETEER 120
QY 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
Db 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
QY 181 KRKGFENKIDRTFGIPEDFDYID 205
Db 181 KRKGFENKIDRTFGIPEDFDYID 205

RESULT 2
US-10-087-190-3
; Sequence 3, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary

; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge. Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-3

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60
QY 61 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSOKHASLOKSTIEKAKIGRCETEER 120
Db 61 DCEIGTSNYWAFPSKALHARKHKLKLEVLSESGSOKHASLOKSTIEKAKIGRCETEER 120
QY 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
Db 121 TRLAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
QY 181 KRKGFENKIDRTFGIPEDFDYID 205
Db 181 KRKGFENKIDRTFGIPEDFDYID 205

RESULT 3
US-10-087-190-14
; Sequence 14, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge. Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-14

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLKLEKIAPEKGIITAMSVKEVLSLVDDGMV 60

Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60
Qy 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Qy 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 4

US-10-087-190-20
; Sequence 20, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-20

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60
Qy 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Qy 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 5

US-10-087-190-21
; Sequence 21, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.

; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-21

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60
Qy 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKHASLOKSIKAKIGRCETEER 120
Qy 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWA 180
Qy 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

RESULT 6

US-10-087-190-22
; Sequence 22, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121P1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-22

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGLSABEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLSQSLVDDGMV 60

```

Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
QY 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

```

```

RESULT 7
US-10-087-190-61
; Sequence 61, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-61

```

```

Query Match 100.0%; Score 1047; DB 12; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
QY 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

```

```

RESULT 8
US-10-408-765A-1821
; Sequence 1821, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing

```

```

; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1821
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1821

```

```

Query Match 100.0%; Score 1047; DB 16; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.4e-85;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
QY 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
Db 61 DCEIRIGTSNYWAPPSKALHARKHKLVLSEQLSEGSQKHASLOKSTIEKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
Db 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFEENKIDRTFGIPEDFDYID 205

```

```

RESULT 9
US-10-087-190-44
; Sequence 44, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-44

```

```

Query Match 99.0%; Score 1036.5; DB 12; Length 205;
Best Local Similarity 99.5%; Pred. No. 4.7e-84;
Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLSQSLVDDGMV 60

```

QY 61 DCRIGTSNYWAPPKALHARKHKLVLVLS-OLSEGSQKHASLQKSIKAKIGRCETEE 119
Db 61 DCRIGTSNYWAPPKALHARKHKLVLVLS-OLSEGSQKHASLQKSIKAKIGRCETEE 120
QY 120 RTRAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISW 179
Db 121 RTRAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISW 180
QY 180 AKRFGFEENKIDRTFGIPEDFDYID 205
Db 181 AKRFGFEENKIDRTFGIPEDFDYID 206

RESULT 10

US-10-087-190-45

; Sequence 45, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-45

Query Match 99.0%; Score 1036.5; DB 12; Length 206;
Best Local Similarity 99.5%; Pred. No. 4.7e-84;
Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MSKKGLSABEKRTRMEIFSETKDVFKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Db 1 MSKKGLSABEKRTRMEIFSETKDVFKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
QY 61 DCRIGTSNYWAPPKALHARKHKLVLVLSQ-LSEGSQKHASLQKSIKAKIGRCETEE 119
Db 61 DCRIGTSNYWAPPKALHARKHKLVLVLSQ-LSEGSQKHASLQKSIKAKIGRCETEE 120
QY 120 RTRAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISW 179
Db 121 RTRAKELSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISW 180
QY 180 AKRFGFEENKIDRTFGIPEDFDYID 205
Db 181 AKRFGFEENKIDRTFGIPEDFDYID 206

RESULT 11

US-10-087-190-24

; Sequence 24, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary

; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-24

Query Match 96.6%; Score 1011; DB 12; Length 198;
Best Local Similarity 100.0%; Pred. No. 8.2e-82;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 KGLSAEKRTRMEIFSETKDVFKLEKIAPEKKGITAMSVKEVLQSLVDDGMVDCER 64
Db 1 KGLSAEKRTRMEIFSETKDVFKLEKIAPEKKGITAMSVKEVLQSLVDDGMVDCER 60
QY 65 IGTSNYWAPPKALHARKHKLVLVLSQSLSEGSQKHASLQKSIKAKIGRCETERTLA 124
Db 61 IGTSNYWAPPKALHARKHKLVLVLSQSLSEGSQKHASLQKSIKAKIGRCETERTLA 120
QY 125 KELSRLDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWAKRKF 184
Db 121 KELSRLDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTNIFAISWAKRKF 180
QY 185 GFEENKIDRTFGIPEDFD 202
Db 181 GFEENKIDRTFGIPEDFD 198

RESULT 12

US-10-087-190-13
; Sequence 13, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-13

Query Match 93.1%; Score 975; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 MMEIFSETKDVFKLEKIAPEKKGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 75

Db 1 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 60
Qy 76 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 135
Db 61 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 120
Qy 136 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 195
Db 121 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 180
Qy 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 13

US-10-087-190-19
; Sequence 19, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-19

Query Match 93.1%; Score 975; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 75
Db 1 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 60
Qy 76 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 135
Db 61 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 120
Qy 136 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 195
Db 121 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 180
Qy 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 14

US-10-087-190-67
; Sequence 67, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.

; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-67

Query Match 93.1%; Score 975; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 75
Db 1 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 60
Qy 76 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 135
Db 61 SKALHARKHKLVELESQSGSQHASLQKSIKAKIGRCETEERTRLAKESSLRDQRE 120
Qy 136 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 195
Db 121 QLKAEVEKYKDCDPQVVEEIRQANKVAKAANRWTDNIFAISWAKRKGFEENKIDRTF 180
Qy 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 15

US-10-087-190-68
; Sequence 68, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-68

Query Match 93.1%; Score 975; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLQSLVDDGMVDCERIGTSNYWAPP 75

Db 1 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDCEIGTSNYIWAPP 60
QY 76 SKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQRE 135
Db 61 SKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQRE 120
QY 136 OLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWAKRKFGFPEENKIDRTF 195
Db 121 OLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWAKRKFGFPEENKIDRTF 180
QY 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 16

US-10-087-190-69
; Sequence 69, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-69

Query Match 93.1%; Score 975; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.2e-78;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDCEIGTSNYIWAPP 75
Db 1 MMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDCEIGTSNYIWAPP 60
QY 76 SKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQRE 135
Db 61 SKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQRE 120
QY 136 OLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWAKRKFGFPEENKIDRTF 195
Db 121 OLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWAKRKFGFPEENKIDRTF 180
QY 196 GIPEDFDYID 205
Db 181 GIPEDFDYID 190

RESULT 17

US-09-799-250-4
; Sequence 4, Application US/09799250
; Publication No. US20030032087A1
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell

; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 121PIF1: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.34US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-799-250-4

Query Match 90.5%; Score 948; DB 10; Length 205;
Best Local Similarity 89.3%; Pred. No. 3.4e-76;
Matches 183; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
QY 1 MSKKKGLSAEEKRTRMMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDV 60
Db 1 MSKKKGLSGEEKRTRMMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDV 60
QY 61 DCEIGTSNYIWAPPSPKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEER 120
Db 61 DCEIGTSNYIWAPPSPKALHARKHKLVLSEQLSEGSQKHASLQKSIKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREOLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 180
Db 121 AMLAKELFSFDQKQKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAISWA 180
QY 181 KRKFGFPEENKIDRTFGIPEDFDYID 205
Db 181 KRKFGFPEENKIDRTFGIPEDFDYID 205

RESULT 18

US-10-087-190-23
; Sequence 23, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-190-23

Query Match 90.5%; Score 948; DB 12; Length 205;
Best Local Similarity 89.3%; Pred. No. 3.4e-76;
Matches 183; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
QY 1 MSKKKGLSAEEKRTRMMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDV 60
Db 1 MSKKKGLSGEEKRTRMMEIFSETKDVFLQKLEKIAPEKIGITAMSVKEVLSLVDGMDV 60

```
QY 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
DB 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
DB 121 AMLAKELSPDRQQLKAEVEKYRECQVVEIRQANKVAKEAANRWTDNIFAISWA 180
QY 181 KRKFGFPEENKIDRTFGIPEDFDYID 205
DB 181 KRKFGFPEENKIDRTFGIPEDFDYID 205

RESULT 19
US-10-087-190-11
; Sequence 11, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-11

Query Match 90.3%; Score 945.5; DB 12; Length 190;
Best Local Similarity 92.2%; Pred. No. 5.1e-76;
Matches 189; Conservative 0; Mismatches 1; Indels 15; Gaps 1;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
DB 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
QY 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
DB 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
DB 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
QY 181 KRKFGFPEENKIDRTFGIPEDFDYID 205
DB 166 KRKFGFPEENKIDRTFGIPEDFDYID 190

RESULT 20
US-10-087-190-18
; Sequence 18, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-11

Query Match 90.3%; Score 945.5; DB 12; Length 190;
Best Local Similarity 92.2%; Pred. No. 5.1e-76;
Matches 189; Conservative 0; Mismatches 1; Indels 15; Gaps 1;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
DB 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
QY 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
DB 61 DCEIRGTSNYWAPPSPKALHARKHKLVLBSQSEGSQKSHASLQKSIKAKIGRCETEER 120
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
DB 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
QY 181 KRKFGFPEENKIDRTFGIPEDFDYID 205
DB 166 KRKFGFPEENKIDRTFGIPEDFDYID 190

RESULT 21
US-10-087-190-62
; Sequence 62, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PIF1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034.20
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-62

Query Match 90.3%; Score 945.5; DB 12; Length 190;
Best Local Similarity 92.2%; Pred. No. 5.1e-76;
Matches 189; Conservative 0; Mismatches 1; Indels 15; Gaps 1;

QY 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
DB 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDDGMV 60
QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
DB 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIFAISWA 180
QY 181 KRKFGFPEENKIDRTFGIPEDFDYID 205
DB 166 KRKFGFPEENKIDRTFGIPEDFDYID 190
```

Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
QY 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
Db 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
QY 121 TRIKELSSLRDQEQKAEVEYKDCDPQVVEIRQANKVAKEAANRWTDNIFAIKSWA 180
Db 121 TRIKELSSLRDQEQKAEVEYKDCDPQVVEIRQANKVAKEAANRWTDNIFAIKSWA 165
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 166 KRKFGFEENKIDRTFGIPEDFDYID 190

RESULT 22
US-10-087-190-63
; Sequence 63, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-63

Query Match 90.3%; Score 945.5; DB 12; Length 190;
Best Local Similarity 92.2%; Pred. No. 5.1e-76;
Matches 189; Conservative 0; Mismatches 1; Indels 15; Gaps 1;
QY 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
QY 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
Db 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
QY 121 TRIKELSSLRDQEQKAEVEYKDCDPQVVEIRQANKVAKEAANRWTDNIFAIKSWA 180
Db 121 TRIKELSSLRDQEQKAEVEYKDCDPQVVEIRQANKVAKEAANRWTDNIFAIKSWA 165
QY 181 KRKFGFEENKIDRTFGIPEDFDYID 205
Db 166 KRKFGFEENKIDRTFGIPEDFDYID 190

RESULT 23
US-10-087-190-53
; Sequence 53, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.

; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-53

Query Match 58.6%; Score 614; DB 12; Length 122;
Best Local Similarity 100.0%; Pred. No. 8.1e-47;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
QY 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
Db 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120
QY 121 TR 122
Db 121 TR 122

RESULT 24
US-10-087-190-9
; Sequence 9, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-9

Query Match 56.4%; Score 591; DB 12; Length 122;
Best Local Similarity 95.9%; Pred. No. 9e-45;
Matches 117; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
Db 1 MSKKKGLSABEKKTRMMEIFSETKDVFKDLKLEKAPKEKGTAMSVKEVLSQSLVDDGMV 60
QY 61 DCRIGTSNYWAPPSPKALHARKHKLKLEVLSESGSQKHSQKSIKAKIGRCETEER 120

```
Db 1 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETAKQ 120
Qy 121 TR 122
Db 121 IK 122

RESULT 25
US-10-087-190-17
; Sequence 17, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-17

Query Match 56.4%; Score 591; DB 12; Length 122;
Best Local Similarity 95.9%; Pred. No. 9e-45;
Matches 117; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Qy 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETEER 120
Db 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETAKQ 120
Qy 121 TR 122
Db 121 IK 122

RESULT 26
US-10-087-190-54
; Sequence 54, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-54

Query Match 56.4%; Score 591; DB 12; Length 122;
Best Local Similarity 95.9%; Pred. No. 9e-45;
Matches 117; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Qy 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETEER 120
Db 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETAKQ 120
Qy 121 TR 122
Db 121 IK 122

RESULT 27
US-10-087-190-55
; Sequence 55, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-55

Query Match 56.4%; Score 591; DB 12; Length 122;
Best Local Similarity 95.9%; Pred. No. 9e-45;
Matches 117; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Db 1 MSKKKGLSAEKKTRMMEIFSETKDVFLQKDLKLEKIAPEKKGITAMSVKEVLQSLVDDGMV 60
Qy 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETEER 120
Db 61 DCEIRGTSNYWAFPPSKALHARKHKLVLVLSQSEGSKHASLQKSIKAKIGRCETAKQ 120
Qy 121 TR 122
Db 121 IK 122

RESULT 28
US-10-087-190-7
; Sequence 7, Application US/10087190
```



```
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-7

Query Match      55.8%; Score 584; DB 12; Length 119;
Best Local Similarity 99.1%; Pred. No. 3.6e-44;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      92 QLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 151
Db      :|||||
        6 ELSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 65

QY      152 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 205
Db      :|||||
        66 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 119

RESULT 29
US-10-087-190-16
; Sequence 16, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-16

Query Match      55.8%; Score 584; DB 12; Length 119;
Best Local Similarity 99.1%; Pred. No. 3.6e-44;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      92 QLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 151
Db      :|||||
        6 ELSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 65

QY      152 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 205
Db      :|||||
        66 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 119

RESULT 29
US-10-087-190-16
; Sequence 16, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-16

Query Match      55.8%; Score 584; DB 12; Length 119;
Best Local Similarity 99.1%; Pred. No. 3.6e-44;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      92 QLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 151
Db      :|||||
        6 ELSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 65

QY      152 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 205
Db      :|||||
        66 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 119

RESULT 30
US-10-087-190-47
; Sequence 47, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034.20
; CURRENT APPLICATION NUMBER: US/10/087,190
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 09/779,250
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-47

Query Match      55.8%; Score 584; DB 12; Length 119;
Best Local Similarity 99.1%; Pred. No. 3.6e-44;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      92 QLSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 151
Db      :|||||
        6 ELSEGSQKHASLQKSIKAKIGRCETEERTRLAKELSSLRDQREQLKAEVEKYKDCDPQV 65

QY      152 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 205
Db      :|||||
        66 VEEIRQANKVAKEAANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPEDFDYID 119

Search completed: September 27, 2004, 09:04:39
Job time : 750 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM_protein - protein search, using sw model

Run on: September 27, 2004, 08:36:32 ; Search time 16 Seconds

(without alignments)
1232.454 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047

Sequence: 1 MSKKKLSAEKTRWEIF.....FEENKIDRTFGIPEDFYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 78:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	349	33.3	210	2 T37610	hypothetical coile
2	197	18.8	128	2 T08972	hypothetical prote
3	168.5	16.1	174	2 S61134	hypothetical prote
4	135.5	12.9	1251	2 JCS368	dynactin 1 - mouse
5	130.5	12.5	1033	2 A41842	dynactin 1 - chicken
6	121.5	11.6	1356	2 S32763	hypothetical prote
7	117	11.2	768	2 T02572	hypothetical prote
8	116.5	11.1	2442	2 T08521	centrosome associa
9	115.5	11.0	880	2 F75103	conserved hypothet
10	114	10.9	284	2 S24972	tropomyosin alpha,
11	114	10.9	764	2 T05409	hypothetical prote
12	113	10.8	189	2 A32183	tropomyosin RPM1 -
13	112	10.7	285	2 JCS199	alpha-tropomyosin
14	112	10.7	285	2 A24199	tropomyosin NM, sk
15	112	10.7	559	2 S49143	Bg10 protein - tap
16	112	10.7	1938	2 A59293	skeletal myosin he
17	111	10.6	559	2 A45620	cyto villin homolog
18	111	10.6	1937	2 I38955	myosin heavy chain
19	110.5	10.6	308	2 T08996	tropomyosin - huma
20	110.5	10.6	629	2 T44607	hypothetical prote
21	110.5	10.6	879	2 C71083	conserved hypothet
22	110	10.5	284	2 I51731	alpha-tropomyosin
23	109.5	10.5	168	2 G86578	CT670 hypothetical
24	109.5	10.5	168	2 B72046	conserved hypothet
25	109.5	10.5	284	2 JCS198	alpha-tropomyosin
26	109.5	10.5	400	2 E70318	hypothetical prote
27	109.5	10.5	1298	2 T24480	hypothetical prote
28	109.5	10.5	1390	2 S51364	sperm tail-specific
29	109	10.4	284	2 JCS2551	tropomyosin alpha

30	109	10.4	284	2 S19691	tropomyosin alpha,
31	109	10.4	676	2 S00084	myosin heavy chain
32	108	10.3	670	2 F84899	hypothetical prote
33	107.5	10.3	1085	2 F96712	hypothetical prote
34	107.5	10.3	1137	2 T19414	hypothetical prote
35	107.5	10.3	1169	2 A64505	Pil5 homolog - Met
36	107	10.2	280	2 A22165	tropomyosin alpha
37	107	10.2	281	2 A34787	tropomyosin 1 alph
38	107	10.2	284	1 TMREA	tropomyosin alpha
39	107	10.2	284	2 B27407	tropomyosin alpha
40	107	10.2	284	2 A39816	tropomyosin 2, fib
41	107	10.2	284	2 A60597	tropomyosin 2, fib
42	107	10.2	284	2 A25825	tropomyosin alpha
43	107	10.2	955	2 S24348	myosin heavy chain
44	106.5	10.2	746	2 T47237	myosin II heavy ch
45	106	10.1	1558	2 B71603	RESA-H3 antigen PF
46	106	10.1	2116	2 A26655	myosin heavy chain
47	105.5	10.1	853	2 T51505	hypothetical prote
48	105	10.0	245	2 C34787	tropomyosin 3 alph
49	105	10.0	248	2 C39816	tropomyosin 5a, fi
50	105	10.0	251	2 B34787	tropomyosin 2 alph
51	105	10.0	284	2 S05445	tropomyosin gamma,
52	105	10.0	791	2 H72552	hypothetical prote
53	104.5	10.0	419	2 G75062	probable flagella-
54	104.5	10.0	705	2 T47949	hypothetical prote
55	104.5	10.0	1177	2 B75150	chromosome segrega
56	104	9.9	248	2 JCS6541	alpha-tropomyosin
57	104	9.9	248	2 I67849	tropomyosin - rat
58	104	9.9	318	2 T49167	hypothetical prote
59	104	9.9	558	2 D70449	conserved hypothet
60	104	9.9	741	2 S39082	myosin heavy chain
61	104	9.9	936	2 S39083	myosin heavy chain
62	104	9.9	1005	2 A64465	hypothetical prote
63	104	9.9	1940	2 A29320	myosin heavy chain
64	104	9.9	2139	2 T18296	myosin heavy chain
65	103.5	9.9	247	2 I67850	tropomyosin - rat
66	103.5	9.9	284	2 S23470	Beta-tropomyosin -
67	103.5	9.9	1061	2 C98690	protein F4H10.4 [
68	103.5	9.9	1822	2 S33441	EF protein - Strept
69	103	9.8	473	2 F70031	cell wall-binding
70	103	9.8	886	2 H69378	conserved hypothet
71	103	9.8	1509	2 A27224	myosin heavy chain
72	103	9.8	1938	1 JX0178	myosin heavy chain
73	103	9.8	3488	2 T34418	hypothetical prote
74	102.5	9.8	472	2 S43554	plasmidogen-bindin
75	102.5	9.8	697	2 T07111	MAR binding filame
76	102	9.7	284	1 TMCHA	tropomyosin 2, ske
77	102	9.7	284	2 A28499	tropomyosin alpha
78	102	9.7	284	2 S24402	alpha-tropomyosin
79	102	9.7	284	2 S66520	tropomyosin, fast
80	102	9.7	284	2 A26113	tropomyosin alpha,
81	102	9.7	372	2 S23326	gene ML2.2 protein
82	102	9.7	476	1 VNU74R	variant surface gl
83	102	9.7	1031	1 A38713	kinesin heavy chai
84	102	9.7	1939	2 T18372	repeat organellar
85	102	9.7	1940	1 S04090	myosin heavy chain
86	101.5	9.7	476	2 S96667	hypothetical prote
87	101.5	9.7	1029	2 H96658	hypothetical prote
88	101.5	9.7	1099	2 G90546	conserved hypothet
89	101.5	9.7	1164	2 T24806	hypothetical prote
90	101.5	9.7	2663	1 S28261	centromere protein
91	101	9.6	226	2 A96268	hypothetical prote
92	101	9.6	952	2 E86147	TiN6.4 protein - A
93	101	9.6	1225	2 A56514	chromokinesin - ch
94	101	9.6	1426	2 T00337	hypothetical prote
95	100.5	9.6	284	1 TMCHS2	tropomyosin 2, giz
96	100	9.6	671	2 H64503	hypothetical prote
97	100	9.6	1017	2 PC4035	cell-cycle-depende
98	100	9.6	1269	2 F84730	probable myosin he
99	100	9.6	1549	2 T21809	hypothetical prote
100	99.5	9.5	221	2 S05446	tropomyosin, nonmu

ALIGNMENTS

RESULT 1

T37610
hypothetical coiled-coil protein - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37610
R:Runt, S.; Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, September 1995
A:Reference number: Z21730
A:Accession: T37610
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residue type: 210
A:Cross-references: EMBL:Z54096; PIDN:CAAS0804.1; GSPDB:GN00066; SPDB:SPAC13A11.03
A:Experimental source: strain 972h-; cosmid c13A11
C:Genetics:
A:Gene: SPDB:SPAC13A11.03
A:Map position: 1
A:Introns: 22/3

Query Match	33.3%	Score	349;	DB 2;	Length	210;			
Best Local Similarity	41.6%	Pred. No.	2.7e-16;						
Matches	84;	Conservative	36;	Mismatches	76;	Indels	6;	Gaps	5;

QY	5	KGLSABEKRTRMEIFSETKDVOLFOLKLEKIAPEKGITAMSVKEVLQSLVDVGDGAVDCER	64
DB	4	KGLSLAEKRRRLRALFPHDSKDPFOLKAEVKGSK-KQIVLQTVKDVQLQSLVDVDDNVKTEK	62
QY	65	IGTSNNYWAFFSKALHARKHKLVELESQLESGOKHSLQKSI--BKAKIGRCETE-BRT	121
DB	63	IGTSNNYWSFFSDAKRSGRESVLGSLQALDLDLKQKSKTLDENISPEKSRDNEGTENDAN	122
QY	122	RLAKELLSLRD-QREQLKAEVEKYKCDQCPQVVEIRQANKVAKEAARNWTDNI PAIKSWA	180
DB	123	QYTLSELLHAKESLUKLJKTQUNSLNHCNPFTFLKKNENTKKYWEAANLWTDQIHTLIAPC	182
QY	181	KRKGFEEENKIDRTFGIPEDFD	202
DB	183	-BDMGADTNOIREYCSIPEDLD	203

RESULT 2

T08972
 RESOL1 2
 hypotheical protein F19B15.200 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C/accession: T08972
 R/Sevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, A.
 submitted to the Protein Sequence Database, May 1999
 A/Reference number: Z16519
 A/Accession: T08972
 A/Molecule type: DNA
 A/Residues: 1-128 <BEV>
 A/cross-references: EMBL:AL078470; GSPDB:GN00062; ATSP:F19B15.200
 A/Experimental source: Cultivar Columbia; BAC clone F19B15
 C/genetics:
 A/Gene: ATSP:F19B15.200
 A/Map position: 4
 A/Introns: 13/3; 52/3; 88/3; 109/3

```
Query Match      18.8%; Score 197; DB 2; Length 128;
Best Local Similarity 39.8%; Pred.No.1.7e-06;
Matches         49; Conservative 23; Mismatches 37; Indels 14; Gaps 3;
```

```
QY      16 NMEIPSEKDVQQLDEKIAPEKGITAMSVKEVLQSLVDDGMVDCEIGIFSNTYYWAFP 75
          :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db       1 MQIPIYESQDFLLKELEKMGPK-KGISVSQKDVIQSLVDDDLVNKKIGIS----- 52
          :|::||::||::||::||::||::||::||::||::||::||::||::||:

QY      76 SKALHARKHKLVLESQISEGQSCHASLQKSIEKAKIGRCETEETRLAKELSSLRDQRE 135
          :|::||::||::||::||::||::||::||::||::||::||::||::||:
```

```

Db      53  ---LSRVQK---LESDLQSGNKKRLAELVDOCEALKKGRSESEBTEALQLDIEKKHK 106
Qy      136 QLK 138
      ||
Db      107 DLK 109

RESULT 3
S61134
hypotheical protein YGL183c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypotheical protein G1604
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C;Accession: S61134; S64200
R;Bertani, I.; Coglievina, M.; Zaccaria, P.; Klima, R.; Bruschi, C.V.
submitted to the EMBL Data Library, September 1995
A;Description: The sequence analysis of a 7.9 kb DNA fragment from the left arm
  ive new genes.
A;Reference number: S61128
A;Accession: S61134
A;Molecule type: DNA
A;Residues: 1-174 <ER>
A;Cross-references: EMBL:X91489; NID:g1143557; PIDN:CAA62791.1; PID:e199057; P:
R;Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64183
A;Accession: S64200
A;Molecule type: DNA
A;Residues: 1-174 <ERU>
A;Cross-references: EMBL:272705; NID:g1322796; PIDN:CAA96895.1; PID:e243495; P:
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:MND1
A;Cross-references: SGD:S0003151
A;Map position: 7L

```

```
Query Match      16.1%; Score 168.5; DB 2; Length 174;
Best Local Similarity 27.2%; Pred. No. 0.00018;
Matches 46; Conservative 42; Mismatches 68; Indels 13; Gaps 5;

Qy    45 MSVKEVLQSIVD-DGMVDCRIGTSNYYWAFPSKALHARKHKLEVLSEQLSEGSKQHSL 103
       |||:::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db     1 MTVKDLVQQMIDEGVISVEKCGNINITYWCFNQTLQKLYDSSELKKIOEVKCDIATY 60
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy   104 QXSIEKA-KIGRCETEERTLAKESLSLRDREOLKAEEVKYKDCDPQVVEIR----- 156
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db    61 KQELDTLTATGRKKFTVGOKSYNREALLEKRKKIQDEIKK-KSNLSLKQISRWDAAKI 119
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   157 QANK----VAKEAANRTWDNIFAIKSWAKEKTFGEENKIDRTGTGIPEDF 201
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db   120 QSNKOQIEPLKVHVLEKTDDNIETIIIDLYVKKFKPEOIRKEFGPEEF 168
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
```

RESULT 4
JC5368
dynactin 1 - mouse
N;Alternate names: p150 Glued
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 05-Nov-1999
C;Accession: JC5368
R;Jang, W.; Weber, J.S.; Tokito, M.K.; Holzbaaur, E.L.F.; Meisler, M.H.
Biochem. Biophys. Res. Commun. 231, 344-347, 1997
A;Title: Mouse p150Glued (dynactin 1) cDNA sequence and evaluation as a candidate for the
A;Reference number: JC5368; MUID:97223454; PMID:9070275
A;Accession: JC5368
A;Molecule type: mRNA
A;Residues: 1-1281 <JAN>
A;Cross-references: GB:U60312; NID:g2104494; PIDN:AAB5773.1; PID:g2104495
A;Experimental source: brain
C;Comment: This protein is a member of the oligomeric dynactin complex that is required f
C;Genetics:
A;Gene: Dctn1
A;Map position: 6

```
Query Match      12.9%; Score 135.5; DB 2; Length 1281;
Best Local Similarity 23.4%; Pred. No. 0.24;
Matches 50; Conservative 43; Mismatches 106; Indels 15; Gaps 5;

QY 2 SKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAMSVKEVLQSLVDGMDV 61
Db 279 ARKEAKEALAKERYMEADTADALEMATDKEMAEERAEISLQOEVEALKERVDLTDD 338
QY 62 CERI-----GTSNYWYAPPSKALHAR-KHKLEVLSEQLSEGSQKSHASLOKSIEKA 110
Db 339 LEILKABIEEKGSDGAASSYQLQLEQONARLKDALVRMDLSLSEKQEHVKLQKLEK- 397
QY 111 KIGRCET--FEETRLAKELSSLDREQLKAEVKEVKDCDPQVVEIRQANKVAKAANR 168
Db 398 KNOELVVROORERLOELSOASTIDELKEQVDALGAE-ENVEMLTORNLMLEEKVRE 456
QY 169 WTONIFAIKSWAKRKFFENKIDRTFGIPEDFD 202
Db 457 LREIVGDLAEMNMNDLQENARETELELREQLD 490

RESULT 5
A41642
C:Species: Gallus gallus (chicken)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 10-Sep-1997
C:Accession: A41642
R:Gill, S.R.; Schroer, T.A.; Szilak, I.; Steuer, E.R.; Sheetz, M.P.; Cleveland, D.W.
J. Cell Biol. 115, 1639-1650, 1991
A:Title: Dynactin, a conserved, ubiquitously expressed component of an activator of vesicle transport
A:Reference number: A41642; MUID:92098576; PMID:1836789
A:Accession: A41642
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1053 <GIL>
A:Cross-references: GB:X62773; NID:G63920; PID:G63921
C:Keywords: cytoskeleton

Query Match      12.5%; Score 130.5; DB 2; Length 1053;
Best Local Similarity 23.6%; Pred. No. 0.41;
Matches 53; Conservative 42; Mismatches 93; Indels 37; Gaps 7;

QY 2 SKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITAM-----SVKEVLQSLVD 57
Db 101 AKKEAKDALBAKERYMEADTADALEMATLDEMAEERAEISLQOEVDLSLKEVYLTD 160
QY 58 GMV---DCERIGTSNYWYAPPSKALHARKHLE---VLESLQSEG-SQKSHASLOKSIEKA 110
Db 161 LEILKHEIEEKGSDGAASSYQVQLBEQONARLKDALVRMDLSASEKQBHVKLQKQMEKK 220
QY 111 KIGRCETEERTRIAKELSSLDREQLKAEVEK-YKDCD-----POVVEIRQ 157
Db 221 N-----TELESRLQQRKLEQVEVKQAEKTVDELKEQVDAALGAEMVETLRE 267
QY 158 ANKVAKEAANRWTDNTPAISKWAKRKFFENKIDRTFGIPEDFD 202
Db 268 RNLDEKVKRELRTVGDLEAMNEMNDELQENARETELELREQLD 312

RESULT 6
S32763
kinectin 1 - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S32763; I37947
R:Kruppa, G.; Fuetterer, A.; Lemke, H.; Kroenke, M.
submitted to the EMBL Data Library, April 1993
A:Description: Cloning and characterization of TAF, a novel transactivating protein.
A:Reference number: S32763
A:Accession: S32763
A:Status: preliminary
A:Molecule type: DNA
```

```
A:Residues: 1-1356 <KRU>
A:Cross-references: EMBL:Z22551
R:Fuetterer, A.; Kruppa, G.; Kramer, B.; Lemke, H.; Kronke, M.
Mol. Biol. Cell 6, 161-170, 1995
A:Title: Molecular cloning and characterization of human kinectin.
A:Reference number: I37947; MUID:95306853; PMID:7787243
A:Accession: I37947
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1356 <RES>
A:Cross-references: EMBL:Z22551; NID:G296163; PIDN:CAA80271.1; PID:G296164
C:Genetics:
A:Gens: GDB:KTN1
A:Cross-references: GDB:6165852; ONIM:600653

Query Match      11.6%; Score 121.5; DB 2; Length 1356;
Best Local Similarity 22.7%; Pred. No. 2.1;
Matches 50; Conservative 31; Mismatches 88; Indels 51; Gaps 5;

QY 3 KKKGLSAEKKTRMMEIFSETKDVFLQKLEKIAPKEKGITA--MSVKEVLQSLVDGMDV 60
Db 1025 RKQNDLRKQWEAMEALASTEKMLQDKNKTSKERQQQVEAVELEAKEVLUKLPFKVSV 1084
QY 61 DCERIGTSNYWYAPPSKALH-----ARKHLEVLSEQLSEGSQKSHASLOKSIEKAKIGRC 115
Db 1085 P-SNLSYGEWLHGFPEKAKCEWAGTSGSEVVKVLEHKLKAEDEMTLLQLECEKYKSVLA 1143
QY 116 ETE-----ERTELA-KELSSLRD 132
Db 1144 ETEGILQKLSVGEQENKWKVDESHKTIKQWSSFTSSEQLERLSENKOIENLR 1203
QY 133 QREQLKAEVKEVKDCDPQVVEIRQANKVAKAANRWTDN 172
Db 1204 EREHLEMELEKAEEMERSTYVTEVRELKDLLELQKLUDDS 1243

RESULT 7
T02572
hypothetical protein At2g39300 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T16B24.6
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02572; F84815
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.;
submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.
A:Reference number: Z14679
A:Accession: T02572
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-768 <ROU>
A:Cross-references: EMBL:AC004697; NID:G3402671; PID:G3402677
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84815
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-768 <STO>
A:Cross-references: GB:A8002093; NID:G3402677; PIDN:AAC28980.1; GSPDB:GN00139
C:Genetics:
A:Gene: T16B24.6; At2g39300
A:Map position: 2
A:Introns: 80/2; 665/3

Query Match      11.2%; Score 117; DB 2; Length 768;
Best Local Similarity 22.1%; Pred. No. 2.3;
Matches 50; Conservative 46; Mismatches 90; Indels 40; Gaps 9;
```

QY 1 MSKKKGLSAB-----EKRTMMEIFSETK-----DVFQL-KOLEKIAPKEKGIT 43
Db 247 MCKEDDVSSLEKRYKEAEKRVKLLSEMEBEKFLSDCPDIFSSLVGDIQOMEERVGL- 305
QY 44 AMSKVEVLQSLVDGMDVDCERIGTSNYWAFPSKALHARKHLEV-LESQSEGSKHAS 102
Db 306 AFEVLSLRQMDERASTREDIRRVKNDWLLKRLKEKTELOVQLETELDRSEWTS 365
QY 103 LOKSIEKAKIGRCETEERTLAKELSLRDQRLKAUEVKYKDCDPQVVEIRQANKVA 162
Db 366 ---KVSEFKV-----EKKRLRVRVLAHNVLSQREISTHEKETERIDMIRHLDET 416
QY 163 KE-----AANRWTDNIPAIKSWAKRKFGFENKIDRTFGIPEDFYI 204
Db 417 TELSATAEEMREENFLMUNLSKLQESYT-----GSTDDLDDYV 454

RESULT 8
T08621
centrosome associated protein CEP250 - human
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08621
R:MacK, G.J.; Rees, J.; Sandblom, O.; Balczon, R.; Fritzel, M.J.; Rattner, J.B.
A:Title: Autoantibodies to a group of centrosomal proteins in human autoimmune sera read
A:Reference number: Z16462; MUID:98165428; PMID:9506584
A:Accession: T08621
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-2442 <MAC>
A:Cross-references: EMBL:AF022655; NID:G28332236; PIDN:AAC06349.1; PID:G2832237
A:Experimental source: cell line HeLa

Query Match 11.4%; Score 116.5; DB 2; Length 2442;
Best Local Similarity 19.7%; Pred. No. 8.6;
Matches 46; Conservative 47; Mismatches 64; Indels 77; Gaps 8;

QY 1 MSKKKGLSABEKRTMMEIFSETKDVFLQKDLK- -KIAPKEKITAMSKVEVLQSLVDG 58
Db 1662 LQKRIQVLEQDRTQTKILEE-----DLEQIKLSUREGRELTTORQLMQERAEG 1713
QY 59 M-----VQCE-----RIGTSNYW 72
Db 1714 KPSPKAQSGSLHMKLILRDKEREVEQCOBHIHELQELQDLQQLGLHRKVGET---- 1769
QY 73 AFPSKALHARKHLEVLSQSLSE-----GSKHASLQKSLKAKIGRCETEERTLAKELS 128
Db 1770 ---SLLSQREQEIVLVQQLQEARQOGELKEQSLSQSLQDEAORALAQRDQ-----ELE 1820
QY 129 SLRDQRLKAUEVKYKDCDPQVVEIRQANKVAKEAANRWTDNIPAIKSWAKR 182
Db 1821 ALQOEQQQAQOEKRVKEKADALQGALEQAHTLKERHGLQDH----KEQARR 1870

RESULT 9
F75103
conserved hypothetical protein FAB0812 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 02-Feb-2001
C:Accession: F75103
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75103
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-880 <KAW>
A:Cross-references: GB:AJ248286; GB:AL096836; NID:G5458366; PIDN:CAB50131.1; PID:G545864
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: FAB0812

C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 11.0%; Score 115.5; DB 2; Length 880;
Best Local Similarity 24.1%; Pred. No. 3.3;
Matches 58; Conservative 38; Mismatches 78; Indels 67; Gaps 10;

QY 3 KKKGL-----SABEKTRMMEIFSETKDVFLQKDLKLEKIAPKEKITAMSKVEVLQ 53
Db 257 RKGLLEKIVQIERSIEEKAKISELEEIVKDPQOEKEKRYKLGFR----- 306
QY 54 LVDDGMDVDCERIGTSNYWAFPSKAL-----HARKHLEVLSQSEGSKHASL--- 103
Db 307 --DEYSEKRLRELEKELSKWSELKAIEEVIKEGKKERAEERKLSIEKRLLELKP 364
QY 104 -----QKSIEKAK-----IGRET--EERTLAKELSSL-----RD 132
Db 365 VELEDARQVQKQIERLKAELKGLSPGEVTEKLESLEKTEIEEAKEITTRIGOMEQE 424
QY 133 QRELKA--EVEKYKDCDPQVVEIRQANKVAKEAANRWTDNIPAIKSWAKRKFGFENK 190
Db 425 KNERKKAIEELRKAAGKCPVCGRELTEHK--KELMERYTLEIKKIEBELKRTTE--EERK 481
QY 191 I 191
Db 482 L 482

RESULT 10
S24972
tropomyosin alpha, cardiac - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 13-Aug-1999
C:Accession: S24972
R:Whitby, F.G.; Kent, H.M.; Stewart, F.; Stewart, M.; Xie, X.; Hatch, V.; Cohen, C.; Phil
submitted to the EMBL Data Library, April 1992
A:Description: Structure of tropomyosin at 9 Angstroms resolution.
A:Reference number: S24972
A:Accession: S24972
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-284 <WHI>
A:Cross-references: EMBL:X66274; NID:G1926; PIDN:CAA46986.1; PID:G1927
C:Superfamily: tropomyosin
C:Keywords: cardiac muscle; heart

Query Match 10.9%; Score 114; DB 2; Length 284;
Best Local Similarity 23.2%; Pred. No. 1.2;
Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;

QY 1 MSKKKGLSAB---EKRTMMEIFSETKDVFLQKDLKLEKIAPKEKITAMSKVEVLQSLVD 57
Db 73 LAEKATDAEADVASLNRIQLFEEELDRAQ-----ERLA-----TALQKLEAEKAADE 122
QY 58 ---GWDCEIGTSNYWAFPSKALHAR---KHKLEVLSQSLSEGSQKH----- 100
Db 123 SERGM-----KVIESRAQKDEBEKMEIQIQLKEA--KHIAEDADSKYE 163
QY 101 -----ASLOKSTIEKAKI--GRCETEERTLAKELSLRDQRLKAUEVKYKDCD 148
Db 164 BVARKLVIESDLERAERAEALSEGK-----ALEEELKTVTNLNKSLAEQAERKYSQKE 218
QY 149 PQVVEIRQANKVAKEAANRWTDNIPAIKSWAKRKFGFENKID 192
Db 219 DRYEEIKVLSDKLEAEETR-----AEFAERSVTKLEKSID 254

RESULT 11
T05409
hypothetical protein F10M6.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05409
R:Bevan, M.; Weichselgartner, M.; Fartmann, B.; Granderath, K.; Dauner, D.; Herzi, A.; Ne

submitted to the Protein Sequence Database, February 1998

A;Reference number: Z15414
A;Accession: T05409
A;Molecule type: DNA
A;Residues: 1-764 <BEV>
A;Cross-references: EMBL:AL021811
A;Experimental source: cultivar Columbia; BAC clone F10M6
C;Genetics:
A;Map position: 4
A;Note: F10M6.170

Query Match	10.9%;	Score 114;	DB 2;	Length 764;
Best Local Similarity	22.8%;	Pred. No. 3.5;		
Matches	42;	Conservative 40;	Mismatches 78;	Indels 24; Gaps 5;
QY	8	SABEKTWMIEIFGETKDVFKOLEKIAPEK-----GITAMSVKEVLSQIV---DDG	58	
Db	49	SASANVLFDKLFARTHLERQTNQHSVYPPDDDLFYSNLGVLESDEALVALVALLKEED	108	
QY	59	MVDCERIGTNNYWFPSKALHARKHLEVLSOLSEGSQKHASLOKSIEKAKIGRCETE	118	
Db	109	LHDAERKLLSD-----KNKLNRAKEBELEKREKTISEASLKHESLOBELKNAV---ELA	159	
QY	119	ERTFLAKELSSLRDRQELKAEVEKYKDCDPQVVEETRQANKVAKAEANWTDNIFAKS	178	
Db	160	SQAREIEELKHKLRRDERDEARAALQSSITLKEELEKTRQ-----EIANRSKEVSMALISE	213	
QY	179	WAKR	182	
Db	214	FESK	217	

RESULT 12

A32183
tropomyosin TPm1 - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein N2312; protein YNL079c
C/Species: *Saccharomyces cerevisiae*
C/Date: 08-Sep-1989 #sequence revision 08-Sep-1989 #text_change 20-Jun-2000
C/Accession: A32183; S53899, S63011; S63018; S63928
R/Liu, H.; Bretschger, A.
Cell 57, 233-242, 1989
A/Title: Disruption of the single tropomyosin gene in yeast results in the disappearance
A/Reference number: A32183; MUID:89195234; PMID:2649250
A/Accession: A32183
A/Molecule type: DNA
A/Residues: 1-199 <LIU>
A/Cross-references: EMBL:M25501; NID:gl73037; PIDN:AAA35174.1; PID:gl73038
R/Poehlmann, R.; Philippsen, P.
submitted to the EMBL Data Library, April 1995
A/Reference number: S53896
A/Accession: S53899
A/Molecule type: DNA
A/Residues: 1-199 <POE>
A/Cross-references: EMBL:X86470; NID:g791101; PIDN:CAA60179.1; PID:g791105
R/Poehlmann, R.; Philippsen, P.
submitted to the Protein Sequence Database, April 1996
A/Reference number: S62997
A/Accession: S63011
A/Molecule type: DNA
A/Residues: 1-199 <POW>
A/Cross-references: EMBL:Z71355; NID:gl301970; PIDN:CAA95953.1; PID:gl301971; MIPS:YNL070
A/Experimental source: strain S288C
R/Soler-Mara, A.; Saiz, J.E.; Ballesta, J.P.G.; Remacha, M.
submitted to the Protein Sequence Database, April 1996
A/Reference number: S63018
A/Accession: S63018
A/Molecule type: DNA
A/Residues: 1-199 <SQL>
A/Cross-references: EMBL:Z71355; NID:gl301970; PIDN:CAA95953.1; PID:gl301971; MIPS:YNL070
A/Experimental source: strain S288C
R/Poehlmann, R.; Philippsen, P.
Yeast 12, 391-402, 1996
A/Title: Sequencing a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV reveals 12

A;Reference number: S63925; MUID:96267764; PMID:8701611
A;Accession: S63928
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-199 <PDF>
A;Cross-references: EMBL:X86470; NID:g791101; PIDN:CAA60179.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Li
C;Genetics:
A;Gene: SGD:TPM1
A;Cross-references: SGD:S0005023; MIPS:YNL079C
A;Map position: 14L
C;Superfamily: tropomyosin TPM1
C;Keywords: coiled coil; cytoskeleton

Query Watch	10.8%;	Score 113;	DB 2;	Length 199;
Best Local Similarity	24.7%;	Prod No. 0.94;		
Matches 46;	Conservative 40;	Mismatches 54;	Indels 45;	Gaps 8
Qy	7	LSAEKPTRRMWEIFSEYKDYFQLKLEKIAPKEKGTAMSVKEVLSQVDD-----GMV	60	
Db	13	LEASWQKYEELKEKNDLEQ--ENVE---KENQIKSLVTVK--QQLEDETEKLEAGLS	65	
Qy	61	DCERIGTSNYVWAPPSKALHARKHL-----EVLSQSLSEG-----SOK	99	
Db	66	DSKQTEQDNVSEKNQIKSLVTYKQHLEEEIEKLEAEALAESKOLSEDSHHLSQNNDFSKK	125	
Qy	100	HASLQKSTIEKAKIGRCETEERTR-----LAKELSSLRDQREQKAEVE-----KYXDC	147	
Db	126	NQOOLEDEESDITKLKETTTEKLRESDLKADQLERRVAALAEQREWERKNEELTVTKYEDA	185	
Qy	148	DPQVVE	153	
Db	186	KKELDE	191	

RESULT 13

JC6199
 alpha-tropomyosin S-1 - axolotl
 C:Species: Ambystoma mexicanum (axolotl)
 C:Date: 11-Aug-1997 #sequence_revision 09-May-1997 #text_change 13-Aug-1999
 C:Accession: JC6199
 R:Yuque, E.A.; Spinner, B.J.; Dube, S.; Dube, D.K.; Lemanski, L.F.
 Gene 185 175-190 1997
 A:Title: Differential expression of a novel isoform of alpha-tropomyosin in
 A:Reference number: JC6199; MCLB:97208870; PMID:9055812
 A:Contents: skeletal muscle
 A:Accession: JC6199
 A:Molecule type: mRNA
 A:Residues: 1-284 <LUQ>
 C:Cross-references: GS:U33450; NID:G1871357; PID:G1871358
 C:Comment: This protein is a actin-binding protein.
 C:Genetics:
 A:Gene: ATMs-1
 C:Superfamily: tropomyosin
 C:Keywords: actin binding

Query Match	10.7%;	Score 112;	DB 2;	Length 284;	
Best Local Similarity	22.7%;	Pred. No. 1.6;			
Matches	48;	Conservative 41;	Mismatches 74;	Indels 48;	Gaps 9;
QY	1	MSKKGLSABE	-----KPTRMVEIFSETKDVFQ	-----LKDLEKIAPK-EKGITAMSV	47
Db	73	LARKATDAESDVASLNR	QLVVEELDRAQRLATALQKLEAEAKADESERGMKV	EN 132	
QY	48	KEVLQSLVDDGMVDCERIGTS	NYWAPPKALHARKHKLEVL	ESQLSEGSQK	---HASL 103
Db	133	R-----ALKDEEKMELOEI	-----QLQEAKHIAEEADRKYEEVARKLV	IEGDL 176	
QY	104	QKSTIEKAKI	--CRCTEETRLAKELISLLEDDQEQ	LEKAEVKKYKCDQPVVVEELRQANKV	161
Db	177	FRABERAEISGKC	-----AELEEEELKVTNNILKSLEAQAEKYSQKEDYBEEIKVL	TDK 231	
QY	162	AKEANRWTDNIFAIKSWARKKFGF	ENKID 192		

Db 232 LKEATR-----AFFAERTVAKLEKSID 254

RESULT 14

A24199

tropomyosin NM, skeletal muscle - human

C:Species: Homo sapiens (man)

C>Date: 02-Jun-1998 #sequence_revision 02-Jun-1998 #text_change 13-Aug-1999

C:Accession: S06210; A24199

R:Clayton, L.; Reinach, F.C.; Chumbley, G.M.; MacLeod, A.R.

J. Mol. Biol. 201, 507-515, 1988

A:Title: Organization of the hTM(nm) gene. Implications for the evolution of muscle and

A:Reference number: S02554; MUID:88332987; PMID:3418707

A:Accession: S06210

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-285 <CIA>

R:Reinach, F.C.; MacLeod, A.R.

Nature 322, 648-650, 1986

A:Title: tissue-specific expression of the human tropomyosin gene involved in the genera

A:Reference number: A24199; MUID:86311274; PMID:3018581

A:Accession: A24199

A:Molecule type: mRNA

A:Residues: 1-285 <REI>

A:Cross-references: GB:X04201; NID:G37429; PIDN:CAA27798.1; PID:G37430

A>Note: an intronless pseudogene resembling this mRNA is also known

C:Superfamily: tropomyosin

C:Keywords: alternative splicing; coiled coil; muscle; skeletal muscle

Query Match 10.7%; Score 112; DB 2; Length 285;

Best Local Similarity 23.5%; Pred. No. 1.6;

Matches 52; Conservative 41; Mismatches 72; Indels 56; Gaps 9;

QY 1 MSKKKGLSABEK---RTRMMEIFSETKDVFPQ-----LKDLKTIAPK-EKGITAMSV 47

DB 74 LAEKQADAEAFVAVSLNRRQLVEELDRQAERLATALOKLEAEAEKAESEKGMKVIEN 133

QY 48 KEVLQSLVDGMDVCERIGTSNYTWAPPFKALHARKHKLVELESQSEGSQKSHASLQKSI 107

DB 134 R---ALKDSEKVELQEI-----QLKAEKHIAEADRKVEEVARKLVIIIEGDL 177

QY 108 EKAKIGRCETETETRLA-----KELSLRQRLQKAEVEKYKDCDPQVVEIRQAN 159

DB 178 ER-----TEERAEALAEKSCSELEBELKNVTNNLSLEAQAEKYSQKEDKYEIEKILT 230

QY 160 KVAKAEANRWTDNIFAIKSWAKKFGFEENKIDRTFGIPED 200

DB 231 DKLEAEATRAE---FAERSVA-----KLEKTIIDDL 259

RESULT 15

S49143

EG10 protein - tapeworm (Echinococcus granulosus)

C:Species: Echinococcus granulosus

C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 24-Nov-2003

C:Accession: S49143

R:Frosch, P.M.; Hartmann, M.; Sygulla, L.; Margutti, P.; Frosch, M.

submitted to the EMBL Data Library, January 1994

A:Description: Identification of a cDNA clone from the larval stage of Echinococcus gran

A:Reference number: S49143

A:Accession: S49143

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-559 <FRO>

A:Cross-references: EMBL:Z29489; NID:G509759; PIDN:CAA82625.1; PID:G509760

C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology

F:12-293/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match 10.7%; Score 112; DB 2; Length 559;

Best Local Similarity 27.8%; Pred. No. 3.4; Indels 32; Gaps 7;

Matches 49; Conservative 34; Mismatches 61; Indels 32; Gaps 7;

QY 1 MSKKKGLSABEKTRMMEIFSETKDVFPQLDKLEKIAPKEKGITAMSVKEVLQSLVDGMDV 60

DB 294 MRRKSDSIEVQWKI-----QAKERELEAEARQLKEERLQRMENEQKRLRLAQ-MV 347

QY 61 DCEIRIGTSNYTWAPPFKALHARKHKLVELESQSE-----GSQKHA--SLQKSIEK-AKI 112

DB 348 EKE-----SDLADMKNKASAYESKIAELMLLQQRHARESLOKSDKLAEM 394

QY 113 GRCETETETRLAKELSLRQRLQKAEVEKYKDCDPQVVEIRQANQKVAKEANR 168

DB 395 NRKLKEETAASAEERNLMAQRDEVOQEVEAQR-----VNAKKEAEKAKQAEEELR 445

RESULT 16

A59293

skeletal myosin heavy chain - domestic rabbit

C:Species: Oryctolagus cuniculus

C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000

C:Accession: A59293

R:Maeda, K.; Hostinova, E.; Roesc.Klein Kauf, A.; Schuster, H.; Gasperik, J.; Wittinghofel

submitted to GenBank, July 1995

A:Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal musc

A:Reference number: A59293

A:Accession: A59293

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1938 <MAE>

A:Cross-references: GB:U32574; NID:G940232; PIDN:AAA74199.1; PID:G940233

A:Experimental source: strain New Zealand White; cell type skeletal muscle fiber type III

C:Genetics:

A:Gene: MHC

C:Superfamily: myosin heavy chain; myosin motor domain homology

F:89-769/Domain: myosin motor domain homology <XMO>

Query Match 10.7%; Score 112; DB 2; Length 1938;

Best Local Similarity 22.2%; Pred. No. 13;

Matches 51; Conservative 44; Mismatches 93; Indels 42; Gaps 7;

QY 1 MSKKKGLSABEKTRMMEIFSETKDVFPQLDK--LEKIAPKEKGITAMSV-----VKEV 50

DB 1364 MSKANSEVAQWRRTYETDAIQRTTEELAEAKKLAQRLQDAEEHVEAVNAKACASLEKTQR 1423

QY 51 LQSLVDGMDVCERIGTS-----NYTWAPPFKALHARKHKLVELESQSEGSQKHA 101

DB 1424 LQNEVEDLMDIVETNACAAADKKQNF-----DKLAEWKHYETHAELEASQKESR 1478

QY 102 SLQKSIEKAK-----IGRCET--EETRLAKELSLRQ-----REQLKAEVEK 143

DB 1479 SLSTEVFKVKNAYEESLDQLETLKRENKLNQOEISDLTEQIAEGKRIHELEKYKQVEQ 1538

QY 144 YKDCDPQVVEIRQANQKVAKEANRWTDNIFAIKSWAKKFGFEENKIDR 193

DB 1539 EKSELQAALAEAEASLEHEGKILRIQLELNQVAKSEIDRKIAEKDEEIDQ 1588

RESULT 17

A45620

Cyovollin homolog - tapeworm (Echinococcus multilocularis)

C:Species: Echinococcus multilocularis

C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-2003

C:Accession: A45620

R:Frosch, P.M.; Frosch, M.; Pfister, T.; Schaad, V.; Bitter-Suermann, D.

Mol. Biochem. Parasitol. 48, 121-130, 1991

A:Title: Cloning and characterization of an immunodominant major surface antigen of Echir

A:Reference number: A45620; MUID:92107215; PMID:1762625

A:Accession: A45620

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-559 <FRO>

A:Cross-references: GB:M61186; NID:G158858; PIDN:AAA29063.1; PID:G158859

A>Note: sequence extracted from NCBI backbone (NCBI:74858, NCBI:74860)

C:Superfamily: ezrin/radixin/moesin family; protein 4.1 membrane-binding domain homology

F:12-293/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match 10.6%; Score 111; DB 2; Length 559;
Best Local Similarity 27.8%; Pred. No. 4;
Matches 49; Conservative 33; Mismatches 62; Indels 32; Gaps 7;
QY 1 MSKKGLSABEKTRMMEIFSETKDFQKDLKIAPEKKGITAMSVKEVLSQSLVDGMV 60
DB 294 MRRKSDSIEVQMKI-----QAKEERELAEERQLKEERLQRMENEQKLRELRAQ-WV 347
QY 61 DCEGIGTSNYWAPPSPKALHARKHKLVLSELSQSE-----GSKGHA--SLQKSIK-AKI 112
DB 348 EKE-----SDLADMKNKASAYESKIAELMLLQQRHARESLQKSDKLAE 394
QY 113 GRCETEERTRIAKELSLRQREOLKAEVEKYKDCDPQVVEEIRQANKVAKAEANR 168
DB 395 NRKLKEETASAEERDLMAQRDEVEQAEAK-----VAMAKEAKEAKAEAEIR 445
RESULT 18
myosin heavy chain, perinatal skeletal muscle - human
N;Contains: MYOSIN ATPase (EC 3.6.4.1)
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C;Accession: J38055; JH0154; S12459; S09332; A30220; S49478
R;Jullian, E.H.; Kelly, A.M.; Pompidou, A.J.; Hoffman, R.; Schiaffino, S.; Stedman, H.H.
Eur. J. Biochem. 230, 1001-1006, 1995
A;Title: Characterization of a human perinatal myosin heavy-chain transcript.
A;Reference number: J38055; MUID:95324556; PMID:7601129
A;Accession: J38055
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1937 <RES>
A;Cross-references: EMBL:Z38133; NID:G558669; PIDN:CAA86293.1; PID:G558669
R;Karsch-Mizrachi, I.; Feghali, R.; Shows, T.B.; Leinwand, L.A.
Gene 89, 289-294, 1990
A;Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.
A;Reference number: JH0154; MUID:90323631; PMID:2373371
A;Accession: JH0154
A;Molecule type: mRNA
A;Residues: 1-14, 'A', 16-859 <KAR>
A;Cross-references: GB:Y00821
A;Experimental source: skeletal muscle
R;Bober, E.
submitted to the EMBL Data Library, January 1989
A;Reference number: S12459
A;Accession: S12459
A;Molecule type: mRNA
A;Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>
A;Cross-references: EMBL:X51592; NID:G29465; PIDN:CAA35941.1; PID:G29466
A;Experimental source: clone gEMHC-F
R;Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
Eur. J. Biochem. 189, 55-65, 1990
A;Title: Identification of three developmentally controlled isoforms of human myosin heavy chain.
A;Reference number: S09331; MUID:90235862; PMID:1691980
A;Accession: S09332
A;Molecule type: mRNA
A;Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-1376, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-1740, 'NT', 1742-1750, 'X', 1752-1759, 'X', 1761-1768, 'X', 1770-1777, 'X', 1779-1786, 'X', 1788-1795, 'X', 1797-1804, 'X', 1806-1813, 'X', 1815-1822, 'X', 1824-1831, 'X', 1833-1840, 'X', 1842-1849, 'X', 1851-1858, 'X', 1860-1867, 'X', 1869-1876, 'X', 1878-1885, 'X', 1887-1894, 'X', 1896-1903, 'X', 1905-1912, 'X', 1914-1921, 'X', 1923-1930, 'X', 1932-1939, 'X', 1941-1948, 'X', 1950-1957, 'X', 1959-1966, 'X', 1968-1975, 'X', 1977-1984, 'X', 1986-1993, 'X', 1995-2002, 'X', 2004-2011, 'X', 2013-2020, 'X', 2022-2029, 'X', 2031-2038, 'X', 2040-2047, 'X', 2049-2056, 'X', 2058-2065, 'X', 2067-2074, 'X', 2076-2083, 'X', 2085-2092, 'X', 2094-2101, 'X', 2103-2110, 'X', 2112-2119, 'X', 2121-2128, 'X', 2130-2137, 'X', 2139-2146, 'X', 2148-2155, 'X', 2157-2164, 'X', 2166-2173, 'X', 2175-2182, 'X', 2184-2191, 'X', 2193-2200, 'X', 2202-2209, 'X', 2211-2218, 'X', 2220-2227, 'X', 2229-2236, 'X', 2238-2245, 'X', 2247-2254, 'X', 2256-2263, 'X', 2265-2272, 'X', 2274-2281, 'X', 2283-2290, 'X', 2292-2299, 'X', 2301-2308, 'X', 2310-2317, 'X', 2319-2326, 'X', 2328-2335, 'X', 2337-2344, 'X', 2346-2353, 'X', 2355-2362, 'X', 2364-2371, 'X', 2373-2380, 'X', 2382-2389, 'X', 2391-2398, 'X', 2400-2407, 'X', 2409-2416, 'X', 2418-2425, 'X', 2427-2434, 'X', 2436-2443, 'X', 2445-2452, 'X', 2454-2461, 'X', 2463-2470, 'X', 2472-2479, 'X', 2481-2488, 'X', 2490-2497, 'X', 2499-2506, 'X', 2508-2515, 'X', 2517-2524, 'X', 2526-2533, 'X', 2535-2542, 'X', 2544-2551, 'X', 2553-2560, 'X', 2562-2569, 'X', 2571-2578, 'X', 2580-2587, 'X', 2589-2596, 'X', 2598-2605, 'X', 2607-2614, 'X', 2616-2623, 'X', 2625-2632, 'X', 2634-2641, 'X', 2643-2650, 'X', 2652-2659, 'X', 2661-2668, 'X', 2670-2677, 'X', 2679-2686, 'X', 2688-2695, 'X', 2697-2704, 'X', 2706-2713, 'X', 2715-2722, 'X', 2724-2731, 'X', 2733-2740, 'X', 2742-2749, 'X', 2751-2758, 'X', 2760-2767, 'X', 2769-2776, 'X', 2778-2785, 'X', 2787-2794, 'X', 2796-2803, 'X', 2805-2812, 'X', 2814-2821, 'X', 2823-2830, 'X', 2832-2839, 'X', 2841-2848, 'X', 2850-2857, 'X', 2859-2866, 'X', 2868-2875, 'X', 2877-2884, 'X', 2886-2893, 'X', 2895-2902, 'X', 2904-2911, 'X', 2913-2920, 'X', 2922-2929, 'X', 2931-2938, 'X', 2940-2947, 'X', 2949-2956, 'X', 2958-2965, 'X', 2967-2974, 'X', 2976-2983, 'X', 2985-2992, 'X', 2994-3001, 'X', 3003-3010, 'X', 3012-3019, 'X', 3021-3028, 'X', 3030-3037, 'X', 3039-3046, 'X', 3048-3055, 'X', 3057-3064, 'X', 3066-3073, 'X', 3075-3082, 'X', 3084-3091, 'X', 3093-3100, 'X', 3102-3109, 'X', 3111-3118, 'X', 3120-3127, 'X', 3129-3136, 'X', 3138-3145, 'X', 3147-3154, 'X', 3156-3163, 'X', 3165-3172, 'X', 3174-3181, 'X', 3183-3190, 'X', 3192-3199, 'X', 3201-3208, 'X', 3210-3217, 'X', 3219-3226, 'X', 3228-3235, 'X', 3237-3244, 'X', 3246-3253, 'X', 3255-3262, 'X', 3264-3271, 'X', 3273-3280, 'X', 3282-3289, 'X', 3291-3298, 'X', 3299-3306, 'X', 3307-3314, 'X', 3315-3322, 'X', 3323-3330, 'X', 3331-3338, 'X', 3339-3346, 'X', 3347-3354, 'X', 3355-3362, 'X', 3363-3370, 'X', 3371-3378, 'X', 3379-3386, 'X', 3387-3394, 'X', 3395-3402, 'X', 3403-3410, 'X', 3411-3418, 'X', 3419-3426, 'X', 3427-3434, 'X', 3435-3442, 'X', 3443-3450, 'X', 3451-3458, 'X', 3459-3466, 'X', 3467-3474, 'X', 3475-3482, 'X', 3483-3490, 'X', 3491-3498, 'X', 3499-3506, 'X', 3507-3514, 'X', 3515-3522, 'X', 3523-3530, 'X', 3531-3538, 'X', 3539-3546, 'X', 3547-3554, 'X', 3555-3562, 'X', 3563-3570, 'X', 3571-3578, 'X', 3579-3586, 'X', 3587-3594, 'X', 3595-3602, 'X', 3603-3610, 'X', 3611-3618, 'X', 3619-3626, 'X', 3627-3634, 'X', 3635-3642, 'X', 3643-3650, 'X', 3651-3658, 'X', 3659-3666, 'X', 3667-3674, 'X', 3675-3682, 'X', 3683-3690, 'X', 3691-3698, 'X', 3699-3706, 'X', 3707-3714, 'X', 3715-3722, 'X', 3723-3730, 'X', 3731-3738, 'X', 3739-3746, 'X', 3747-3754, 'X', 3755-3762, 'X', 3763-3770, 'X', 3771-3778, 'X', 3779-3786, 'X', 3787-3794, 'X', 3795-3802, 'X', 3803-3810, 'X', 3811-3818, 'X', 3819-3826, 'X', 3827-3834, 'X', 3835-3842, 'X', 3843-3850, 'X', 3851-3858, 'X', 3859-3866, 'X', 3867-3874, 'X', 3875-3882, 'X', 3883-3890, 'X', 3891-3898, 'X', 3899-3906, 'X', 3907-3914, 'X', 3915-3922, 'X', 3923-3930, 'X', 3931-3938, 'X', 3939-3946, 'X', 3947-3954, 'X', 3955-3962, 'X', 3963-3970, 'X', 3971-3978, 'X', 3979-3986, 'X', 3987-3994, 'X', 3995-4002, 'X', 4003-4010, 'X', 4011-4018, 'X', 4019-4026, 'X', 4027-4034, 'X', 4035-4042, 'X', 4043-4050, 'X', 4051-4058, 'X', 4059-4066, 'X', 4067-4074, 'X', 4075-4082, 'X', 4083-4090, 'X', 4091-4098, 'X', 4099-4106, 'X', 4107-4114, 'X', 4115-4122, 'X', 4123-4130, 'X', 4131-4138, 'X', 4139-4146, 'X', 4147-4154, 'X', 4155-4162, 'X', 4163-4170, 'X', 4171-4178, 'X', 4179-4186, 'X', 4187-4194, 'X', 4195-4202, 'X', 4203-4210, 'X', 4211-4218, 'X', 4219-4226, 'X', 4227-4234, 'X', 4235-4242, 'X', 4243-4250, 'X', 4251-4258, 'X', 4259-4266, 'X', 4267-4274, 'X', 4275-4282, 'X', 4283-4290, 'X', 4291-4298, 'X', 4299-4306, 'X', 4307-4314, 'X', 4315-4322, 'X', 4323-4330, 'X', 4331-4338, 'X', 4339-4346, 'X', 4347-4354, 'X', 4355-4362, 'X', 4363-4370, 'X', 4371-4378, 'X', 4379-4386, 'X', 4387-4394, 'X', 4395-4402, 'X', 4403-4410, 'X', 4411-4418, 'X', 4419-4426, 'X', 4427-4434, 'X', 4435-4442, 'X', 4443-4450, 'X', 4451-4458, 'X', 4459-4466, 'X', 4467-4474, 'X', 4475-4482, 'X', 4483-4490, 'X', 4491-4498, 'X', 4499-4506, 'X', 4507-4514, 'X', 4515-4522, 'X', 4523-4530, 'X', 4531-4538, 'X', 4539-4546, 'X', 4547-4554, 'X', 4555-4562, 'X', 4563-4570, 'X', 4571-4578, 'X', 4579-4586, 'X', 4587-4594, 'X', 4595-4602, 'X', 4603-4610, 'X', 4611-4618, 'X', 4619-4626, 'X', 4627-4634, 'X', 4635-4642, 'X', 4643-4650, 'X', 4651-4658, 'X', 4659-4666, 'X', 4667-4674, 'X', 4675-4682, 'X', 4683-4690, 'X', 4691-4698, 'X', 4699-4706, 'X', 4707-4714, 'X', 4715-4722, 'X', 4723-4730, 'X', 4731-4738, 'X', 4739-4746, 'X', 4747-4754, 'X', 4755-4762, 'X', 4763-4770, 'X', 4771-4778, 'X', 4779-4786, 'X', 4787-4794, 'X', 4795-4802, 'X', 4803-4810, 'X', 4811-4818, 'X', 4819-4826, 'X', 4827-4834, 'X', 4835-4842, 'X', 4843-4850, 'X', 4851-4858, 'X', 4859-4866, 'X', 4867-4874, 'X', 4875-4882, 'X', 4883-4890, 'X', 4891-4898, 'X', 4899-4906, 'X', 4907-4914, 'X', 4915-4922, 'X', 4923-4930, 'X', 4931-4938, 'X', 4939-4946, 'X', 4947-4954, 'X', 4955-4962, 'X', 4963-4970, 'X', 4971-4978, 'X', 4979-4986, 'X', 4987-4994, 'X', 4995-5002, 'X', 5003-5010, 'X', 5011-5018, 'X', 5019-5026, 'X', 5027-5034, 'X', 5035-5042, 'X', 5043-5050, 'X', 5051-5058, 'X', 5059-5066, 'X', 5067-5074, 'X', 5075-5082, 'X', 5083-5090, 'X', 5091-5098, 'X', 5099-5106, 'X', 5107-5114, 'X', 5115-5122, 'X', 5123-5130, 'X', 5131-5138, 'X', 5139-5146, 'X', 5147-5154, 'X', 5155-5162, 'X', 5163-5170, 'X', 5171-5178, 'X', 5179-5186, 'X', 5187-5194, 'X', 5195-5202, 'X', 5203-5210, 'X', 5211-5218, 'X', 5219-5226, 'X', 5227-5234, 'X', 5235-5242, 'X', 5243-5250, 'X', 5251-5258, 'X', 5259-5266, 'X', 5267-5274, 'X', 5275-5282, 'X', 5283-5290, 'X', 5291-5298, 'X', 5299-5306, 'X', 5307-5314, 'X', 5315-5322, 'X', 5323-5330, 'X', 5331-5338, 'X', 5339-5346, 'X', 5347-5354, 'X', 5355-5362, 'X', 5363-5370, 'X', 5371-5378, 'X', 5379-5386, 'X', 5387-5394, 'X', 5395-5402, 'X', 5403-5410, 'X', 5411-5418, 'X', 5419-5426, 'X', 5427-5434, 'X', 5435-5442, 'X', 5443-5450, 'X', 5451-5458, 'X', 5459-5466, 'X', 5467-5474, 'X', 5475-5482, 'X', 5483-5490, 'X', 5491-5498, 'X', 5499-5506, 'X', 5507-5514, 'X', 5515-5522, 'X', 5523-5530, 'X', 5531-5538, 'X', 5539-5546, 'X', 5547-5554, 'X', 5555-5562, 'X', 5563-5570, 'X', 5571-5578, 'X', 5579-5586, 'X', 5587-5594, 'X', 5595-5602, 'X', 5603-5610, 'X', 5611-5618, 'X', 5619-5626, 'X', 5627-5634, 'X', 5635-5642, 'X', 5643-5650, 'X', 5651-5658, 'X', 5659-5666, 'X', 5667-5674, 'X', 5675-5682, 'X', 5683-5690, 'X', 5691-5698, 'X', 5699-5706, 'X', 5707-5714, 'X', 5715-5722, 'X', 5723-5730, 'X', 5731-5738, 'X', 5739-5746, 'X', 5747-5754, 'X', 5755-5762, 'X', 5763-5770, 'X', 5771-5778, 'X', 5779-5786, 'X', 5787-5794, 'X', 5795-5802, 'X', 5803-5810, 'X', 5811-5818, 'X', 5819-5826, 'X', 5827-5834, 'X', 5835-5842, 'X', 5843-5850, 'X', 5851-5858, 'X', 5859-5866, 'X', 5867-5874, 'X', 5875-5882, 'X', 5883-5890, 'X', 5891-5898, 'X', 5899-5906, 'X', 5907-5914, 'X', 5915-5922, 'X', 5923-5930, 'X', 5931-5938, 'X', 5939-5946, 'X', 5947-5954, 'X', 5955-5962, 'X', 5963-5970, 'X', 5971-5978, 'X', 5979-5986, 'X', 5987-5994, 'X', 5995-6002, 'X', 6003-6010, 'X', 6011-6018, 'X', 6019-6026, 'X', 6027-6034, 'X', 6035-6042, 'X', 6043-6050, 'X', 6051-6058, 'X', 6059-6066, 'X', 6067-6074, 'X', 6075-6082, 'X', 6083-6090, 'X', 6091-6098, 'X', 6099-6106, 'X', 6107-6114, 'X', 6115-6122, 'X', 6123-6130, 'X', 6131-6138, 'X', 6139-6146, 'X', 6147-6154, 'X', 6155-6162, 'X', 6163-6170, 'X', 6171-6178, 'X', 6179-6186, 'X', 6187-6194, 'X', 6195-6202, 'X', 6203-6210, 'X', 6211-6218, 'X', 6219-6226, 'X', 6227-6234, 'X', 6235-6242, 'X', 6243-6250, 'X', 6251-6258, 'X', 6259-6266, 'X', 6267-6274, 'X', 6275-6282, 'X', 6283-6290, 'X', 6291-6298, 'X', 6299-6306, 'X', 6307-6314, 'X', 6315-6322, 'X', 6323-6330, 'X', 6331-6338, 'X', 6339-6346, 'X', 6347-6354, 'X', 6355-6362, 'X', 6363-6370, 'X', 6371-6378, 'X', 6379-6386, 'X', 6387-6394, 'X', 6395-6402, 'X', 6403-6410, 'X', 6411-6418, 'X', 6419-6426, 'X', 6427-6434, 'X', 6435-6442, 'X', 6443-6450, 'X', 6451-6458, 'X', 6459-6466, 'X', 6467-6474, 'X', 6475-6482, 'X', 6483-6490, 'X', 6491-6498, 'X', 6499-6506, 'X', 6507-6514, 'X', 6515-6522, 'X', 6523-6530, 'X', 6531-6538, 'X', 6539-6546, 'X', 6547-6554, 'X', 6555-6562, 'X', 6563-6570, 'X', 6571-6578, 'X', 6579-6586, 'X', 6587-6594, 'X', 6595-6602, 'X', 6603-6610, 'X', 6611-6618, 'X', 6619-6626, 'X', 6627-6634, 'X', 6635-6642, 'X', 6643-6650, 'X', 6651-6658, 'X', 6659-6666, 'X', 6667-6674, 'X', 6675-6682, 'X', 6683-6690, 'X', 6691-6698, 'X', 6699-6706, 'X', 6707-6714, 'X', 6715-6722, 'X', 6723-6730, 'X', 6731-6738, 'X', 6739-6746, 'X', 6747-6754, 'X', 6755-6762, 'X', 6763-6770, 'X', 6771-6778, 'X', 6779-6786, 'X', 6787-6794, 'X', 6795-6802, 'X', 6803-6810, 'X', 6811-6818, 'X', 6819-6826, 'X', 6827-6834, 'X', 6835-6842, 'X', 6843-6850, 'X', 6851-6858, 'X', 6859-6866, 'X', 6867-6874, 'X', 6875-6882, 'X', 6883-6890, 'X', 6891-6898, 'X', 6899-6906, 'X', 6907-6914, 'X', 6915-6922, 'X', 6923-6930, 'X', 6931-6938, 'X', 6939-6946, 'X', 6947-6954, 'X', 6955-6962, 'X', 6963-6970, 'X', 6971-6978, 'X', 6979-6986, 'X', 6987-6994, 'X', 6995-7002, 'X', 7003-7010, 'X', 7011-7018, 'X', 7019-7026, 'X', 7027-7034, 'X', 7035-7042, 'X', 7043-7050, 'X', 7051-7058, 'X', 7059-7066, 'X', 7067-7074, 'X', 7075-7082, 'X', 7083-7090, 'X', 7091-7098, 'X', 7099-7106, 'X', 7107-7114, 'X', 7115-7122, 'X', 7123-7130, 'X', 7131-7138, 'X', 7139-7146, 'X', 7147-7154, 'X', 7155-7162, 'X', 7163-7170, 'X', 7171-7178, 'X', 7179-7186, 'X', 7187-7194, 'X', 7195-7202, 'X', 7203-7210, 'X', 7211-7218, 'X', 7219-7226, 'X', 7227-7234, 'X', 7235-7242, 'X', 7243-7250, 'X', 7251-7258, 'X', 7259-7266, 'X', 7267-7274, 'X', 7275-7282, 'X', 7283-7290, 'X', 7291-7298, 'X', 7299-7306, 'X', 7307-7314, 'X', 7315-7322, 'X', 7323-7330, 'X', 7331-7338, 'X', 7339-7346, 'X', 7347-7354, 'X', 7355-7362, 'X', 7363-7370, 'X', 7371-7378, 'X', 7379-7386, 'X', 7387-7394, 'X', 7395-7402, 'X', 7403-7410, 'X', 7411-7418, 'X', 7419-7426, 'X', 7427-7434, 'X', 7435-7442, 'X', 7443-7450, 'X', 7451-7458, 'X', 7459-7466, 'X', 7467-7474, 'X', 7475-7482, 'X', 7483-7490, 'X', 7491-7498, 'X', 7499-7506, 'X', 7507-7514, 'X', 7515-7522, 'X', 7523-7530, 'X', 7531-7538, 'X', 7539-7546, 'X', 7547-7554, 'X', 7555-7562, 'X', 7563-7570, 'X', 7571-7578, 'X', 7579-7586, 'X', 7587-7594, 'X', 7595-7602, 'X', 7603-7610, 'X', 7611-7618, 'X', 7619-7626, 'X', 7627-7634, 'X', 7635-7642, 'X', 7643-7650, 'X', 7651-7658, 'X', 7659-7666, 'X', 7667-7674, 'X', 7675-7682, 'X', 7683-7690, 'X', 7691-7698, 'X', 7699-7706, 'X', 7707-7714, 'X', 7715-7722, 'X', 7723-7730, 'X', 7731-7738, 'X', 7739-7746, 'X', 7747-7754, 'X', 7755-7762, 'X', 7763-7770, 'X', 7771-7778, 'X', 7779-7786, 'X', 7787-7794, 'X', 7795-7802, 'X', 7803-7810, 'X', 7811-7818, 'X', 7819-7826, 'X', 7827-7834, 'X', 7835-7842, 'X', 7843-7850, 'X', 7851-7858, 'X', 7859-7866, 'X', 7867-7874, 'X', 7875-7882, 'X', 7883-7890, 'X', 7891-7898, 'X', 7899-7906, 'X', 7907-7914, 'X', 7915-7922, 'X', 7923-7930, 'X', 7931-7938, 'X', 7939-7946, 'X', 7947-7954, 'X', 7955-7962, 'X', 7963-7970, 'X', 7971-7978, 'X', 7979-7986, 'X', 7987-7994, 'X', 7995-8002, 'X', 8003-8010, 'X', 8011-8018, 'X', 8019-8026, 'X', 8027-8034, 'X', 8035-8042, 'X', 8043-8050, 'X', 8051-8058, 'X', 8059-8066, 'X', 8067-8074, 'X', 8075-8082, 'X', 8083-8090, 'X', 8091-8098, 'X', 809

```
RESULT 20
T44607
hypothetical protein hp71 - Halobacterium salinarum
C:Species: Halobacterium salinarum
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T44607
R:Ruepp, A.; Wanner, G.; Soppa, J.
Arch. Microbiol. 169, 1-9, 1998
A:Title: A 71-kDa protein from Halobacterium salinarum belongs to a ubiquitous P-loop A
A:Reference number: 222810; MUID:98060711; PMID:9396829
A:Accession: T44607
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-629 <RUE>
A:Cross-references: EMBL:Y13615; PIDN:CAA73936.1
A:Note: the source is designated as Halobacterium salinarum
C:Genetics:
A:Gene: hp71
C:Function:
A:Description: might be involved in cytoskeleton formation and/or chromosome partitionin
Query Match 10.6%; Score 110.5; DB 2; Length 629;
Best Local Similarity 23.2%; Pred. No. 4.9;
Matches 55; Conservative 49; Mismatches 76; Indels 57; Gaps 11;
QY 2 SKKGLSABEKTRMMIFSETKDVDFOLKLE-----KIAPKEKGITAM 45
DB 215 AKQSDLEREARLERLENQVDRKA-QLDDKEAALSLDIPDSPTAEADIAEKQTRIDEL 273
QY 46 SVKEVL-----QSLVDDG---WVDCRIGTSNYWAFPSKALHARKHKLVELESOL 93
DB 274 AVKIDLDLHRSTKALIDGEGDILITDVERTLSGDTFCFVCGA-----ETTAETVTERL 329
QY 94 SEGQKHSASLQKSTERAKIGRCETETRL---AKELSSLRDQREQLKAEVEKYKDCDPQ 150
DB 330 NEISDRQESLRE---QRATL---TEEVTOQRTRETESKRQQAEBLEDEIKLR-VDIQ 382
QY 151 -----VVEIR---QANKVAKEANRWTDNIPFAKSWAKKFGFEENKIDR 193
DB 383 EDQHEVRSIEATBELQAEIQREAEYEAAGKAGESHSELKTIQKIGSTETKIDR 439
RESULT 21
C71083
conserved hypothetical protein PH0929 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 02-Feb-2001
C:Accession: C71083
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: C71083
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-679 <KAW>
A:Cross-references: GB:AF000004; NID:g3236131; PIDN:BAA30025.1; PID:g3257342
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0929
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032
Query Match 10.6%; Score 110.5; DB 2; Length 879;
Best Local Similarity 22.1%; Pred. No. 7;
Matches 48; Conservative 51; Mismatches 77; Indels 41; Gaps 9;
QY 4 KKGLSABEKTRMMIFSETKDVDF---QLKLEKIAPK--EKGITAMSVKVLQSLVDDG- 58
DB 569 KKELS--EIEDRLRLGFKTIDELSGRILEKFNKYIEAKNAEKRLDILSLXDERE 626
QY 59 -----WVDCRIGT--SNYWAFPSKALHARKH-----LVLESQSLSEGS 97
```

```
DB 627 ELDKAFEBELAKIETDIEKVTSQLNELQKFDQKYEERKMKLSMEIKGLTKLEELE 686
QY 98 QKHASLQKSIKATIGRCETEERFLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQ 157
DB 687 RRRDEIKSTIEKLEBKERESAKMELEKLNIAIKRIEELRGKIKYEKALIKE--EALNK 744
QY 158 ANKVAKEANRWTDNIP---AIKSWAKKFGFEENKI 191
DB 745 IGEIASEIFSETDGKYSIAIRA-----EDNKV 773
RESULT 22
I51731
alpha-tropomyosin - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 13-Mar-1997 #sequence_revision 13-Mar-1997 #text_change 13-Aug-1999
C:Accession: I51731
R:Ohara, O.; Doric, R.L.; Gilbert, W.
Proc. Natl. Acad. Sci. U.S.A. 86, 5673-5677, 1989
A:Title: One-sided polymerase chain reaction: The amplification of cDNA.
A:Reference number: I51731; MUID:89345529; PMID:2788276
A:Accession: I51731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-284 <OHA>
A:Cross-references: GB:M24635; NID:g214973; PIDN:AA50021.1; PID:g214974
C:Superfamily: tropomyosin
Query Match 10.5%; Score 110; DB 2; Length 284;
Best Local Similarity 23.3%; Pred. No. 2.2;
Matches 51; Conservative 44; Mismatches 72; Indels 52; Gaps 10;
QY 1 MSKKGSLASB---EKTRMMEIFSETKDVQ-----LKLEKIAPK-EKGITAMSV 47
DB 73 LAEKATDAEGDVASLNRIQLVEEELDRAQLATLQKLEEAKAADSERGMKVIEN 132
QY 48 KEVLQSLVDDCMVDCRIGTSNYWAFPSKALHARKHKLVELESQSLSEGSQK----HASL 103
DB 133 R-----ALKDEKMBLQEI-----CLKAKHIAEADRKYEAVARKLVIVGEL 176
QY 104 QKSTEKAKI--GRCEETEERFLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKV 161
DB 177 ERTEERAEINEGKC-----SELEELKTVTNMKSLEAQAQKYSAKEDKYEEIKVLTDK 231
QY 162 AKEANRWTDNIPFAKSWAKKFGFEENKIDRTFGIPED 200
DB 232 LKEAETRAE---FAERSVA-----KLEKTIIDDL 258
RESULT 23
G86578
CP670 hypothetical protein [imported] - Chlamydothila pneumoniae (strain J138)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: G86578
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ist
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138...
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: G86578
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-168 <STO>
A:Cross-references: GB:BA000008; NID:98979078; PIDN:BAA98913.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPJ0706
Query Match 10.5%; Score 109.5; DB 2; Length 168;
Best Local Similarity 29.8%; Pred. No. 1.3;
Matches 48; Conservative 29; Mismatches 51; Indels 33; Gaps 8;
```

Qy 10 EKKTRMELFSETKOVFOLEKIAPKEKGITAMSVKEVLSIVDDGKVDCERIGTGN 69
 :|||:::||::||| ::|||::|||:
Dd 25 KERR-BLLETEQE-----KLREKEARDKVNHYMKIOQQ-LRDLDDEGT----- 68
 :|||::|||:||||| |||:|
Qy 70 YYWAFPSKAHLHARKHKLEVLESOLSE-----GSOKIASL--OQSIEKAKIG---RCETE 119
 :|||::|||:||||| |||:|
Dd 69 -----SDAVTLQIKSYIKVVAVQLSREEEKNVKQEVVLAAASKELEKAEVNLAKRKEEE 122
 :|||::|||:||||| |||:|
Qy 120 RTRLAKELSRLDQRQLKAEVEKYDKCDPQVVEETIQANK 160
 :|||::|||:||||| |||:|
Dd 123 KTRLHEEWNMKEALKEEARAE-EKEQDEMGLLFQLRKQKK 162
 :|||::|||:||||| |||:|

RESULT 24

B72046 conserved hypothetical protein CP0040 [imported] - Chlamydophila pneumoniae (strains CWI C; Species: Chlamydophila pneumoniae, Chlamydia pneumoniae C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000 C; Accession: B72046; G81620 R.; Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999 A; Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A; Reference number: A72000; MUID:99206606; PMID:10192388
A; Accession: B72046
A; Molecule type: DNA
A; Residues: 1-168 <ARN>
A; Cross-references: GB:AEO01652; NID:g4376997; PIDN:AAD18845.1; PID:g4377000
A; Experimental source: strain CWLO29
R.; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Winn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A; Reference number: AB1500; MUID:20150255; PMID:10684935
A; Accession: G81620
A; Molecule type: DNA
A; Residues: 1-168 <REA>
A; Cross-references: GS:AEO02167; GB:AEO02163; NID:g7188971; PIDN:AAF37935.1; PID:g718898
A; Experimental source: strain AR39, HL cells
C; Genetics:
A; Gene: CPn0706; CP0040

```

Query Match      10.5%; Score 109.5; DB 2; Length 168;
Best Local Similarity 29.8%; Pred. No. 1.3;
Matches 48; Conservative 29; Mismatches 51; Indels 33; Gaps 8;

QY      10  EKKRTRMWEIPISETKDYFQLKDLEKAPKEGGITAMSVKEVLSQSLVDDGNVDCERIGTSN 69
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      25  KKKR-RLEITEQE-----KLRKEAERDKVKNHYMKIQO-LRDDLDGGTT----- 68

QY      70  YYWAFPSKALFHARKHLEVEESOLSE-----GSOKHASL--QKSIEKAKIG---RCETE 119
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      69  -----SDAVLIQISYIKVAVQLSEEEKVKNKQKEVVLAASKELEKAEVNLAKRKEEE 122

QY      120  RTRLAKELSLRDRQREOLKAEEVKYKDCDPQVVEEIRQANK 160
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      123  KTRLHKEENMKALKEARAE-EKEQDEMGOLLIFQLFQKKEK 162

```

RESULT 25
JC6198
alpha-tropomyosin C-2 - axolotl
C/Species: Ambystoma mexicanum (axolotl)
C/Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 13-Aug-1999
C/Accession: JC6198
R/Luque, E.A.; Spinner, B.J.; Dube, D.K.; Lemanski, L.F.
Gene 185, 175-180, 1997
A/Title: Differential expression of a novel isoform of alpha-tropomyosin in cardiac and
A/Reference number: JC6198; MUID:97208870; PMID:9055812
A/Contents: Heart
A/Accession: JC6198
A/Molecule type: mRNA
A/Residues: 1-284 <LUQ>
A/Cross-references: GB:U33449; NID:q1871355; PIDN:AAC60091.1; PID:q1871356

```
C;Comment: This protein is a actin-binding protein.
C;Genetics:
A;Gene: ATmc-2
C;Superfamily: tropomyosin
C;keywords: actin binding

Query Match          10.5%; Score 109.5; DB 2; Length 284;
Best Local Similarity 22.6%; Pred.No. 2.4;
Matches 49; Conservative 42; Mismatches 71; Indels 55; Gaps
      9;

QY   2 SKKKGLSAEKR-----TRMEIPISETVDVFQ-----LKLEKIAPK-EKG 41
      |||::|||
Db    67 SEESLLTADAEKAAGSDASLNRRIGQLVEEELDRAQLERLATALQKLEEAKADESERG 126
      ::::|
QY   42 ITAMSVKEVLQSIVDDGVDCERIGTSNYWAFSPKALHARKHKLVLESQSEGSK-- 99
      |||::|||
Db    127 MKVIENR----ALKDEKMKLOEI-----QLQEKAIHAEDRKRYEEVARKLV 170
      :|
QY   100 --HASIQKSIEKAKI--GRCTETERTLAKELSRLDQRDLQAIVEKYKDQPOVVVEI 155
      |||::|||
Db    171 IIEGDLEIAFEERAELSEGGK-----AELEBELTATNNLSLEAQAEKYSQKEDKYBEEI 225
      |||::|||
QY   156 RQANKVAKEAANRWTDNIPAIKSWAKRKFGFEEKID 192
      |||::|||
Db    226 KVLITDKLKEATR-----AEFAERTVAKLEKSID 254
      :|

RESULT 26
E70318
hypothetical protein aa197 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: E70318
R;Deckerdt, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: E70318
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-400 >>>Q>
A;Cross-references: GB:AE000677; NID:g2982900; PIDN:AAC06533.1; PID:g2982915; G
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aaq 197
```

[illegible]

RESULT 27
T24480
hypothetical protein T04H1.4 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text change 17-Mar-2000

```

Db      825 K-----KKCEKAAKKRKEAEKKCKEKTAKKRKETAETKKCEKAAKKRQQAEEKKCEK 878
QY      110 AKIGRCETEERTLA-----KESSLRDQREQLK--AEVEKYKDCDPQVVEEIRQA--N 159
Db      879 AAKRKEAAEKKKCAEAKEKELAEKKCEAAKKEVAERKKCE--ELAKKIKKAAEK 937
QY      160 KVAKAANRWTDNIFAIKSWAKRPGFPEENKIDRTFG 196
Db      938 KKCKKLAK-----KCKAGEKKKLKKKAG 961

RESULT 29
JC2551
tropomyosin alpha chain - axolotl
C:Species: Ambystoma mexicanum (axolotl)
C:Date: 19-May-1995 #sequence_revision 14-Jul-1995 #text_change 13-Aug-1999
C:Accession: JC2551
R:Laque, E.A.; Lemanski, L.F.; Dube, D.K.
Biochem. Biophys. Res. Commun. 203, 319-325, 1994
A:Title: Molecular cloning, sequencing and expression of an isoform of cardiac alpha-tro-
A:Reference number: JC2551; MUID:94354820; PMID:8074673
A:Accession: JC2551
A:Molecule type: mRNA
A:Residues: 1-284 <LUO>
A:Cross-references: GB:L35107; NID:g939982; PIDN:AAA74124.1; PID:g520748
A:Experimental source: heart
C:Superfamily: tropomyosin

Query Match 10.4%; Score 109; DB 2; Length 284;
Best Local Similarity 22.7%; Pred. No. 2.6;
Matches 48; Conservative 40; Mismatches 75; Indels 48; Gaps 9;

QY      1 MSKKKGLSAEE---KRTFMELFSETKDVFQ-----LKDLEKIAPK-EKGITAMSV 47
Db      73 LADKATKAESDVASLNRRIOLVBEELDRAQERLATLQKLEAEKAADSRGKMVIEN 132
QY      48 KEVLQSLVDDGMVDCERIGTSNYWAFPSKALHARKHKLVLVESQLSEGSQK----HASL 103
Db      133 R-----ALNDEEKWELQEI-----QLQEAKHIAEADRKYEVEAVKLVIEGDL 176
QY      104 QKSIFKAKI--GRETEBTRLAKELSLRDQRLQAEVEKYKDCDPQVVEEIRQANKV 161
Db      177 ERAEERAEBSGKC-----AELEBELKTVTNLKSLEAQAQKYSQKEDKYVEEI-KVLTDK 231
QY      162 AKEAANRWTDNIFAIKSWAKRPGFPEENKID 192
Db      232 LKEAETR-----AEFAERTVAKLEKSID 254

RESULT 30
S19691
tropomyosin alpha, skeletal - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Nov-1993 #sequence_revision 19-Oct-1995 #text_change 13-Aug-1999
C:Accession: S19691; S16782
R:Hardy, S.; Fiszman, M.Y.; Osborne, H.B.; Thiebaud, P.
Eur. J. Biochem. 202, 431-440, 1991
A:Title: Characterization of muscle and non muscle Xenopus laevis tropomyosin mRNAs trans-
A:Reference number: S19690; MUID:92104164; PMID:1840524
A:Accession: S19691
A>Status: Preliminary
A:Molecule type: nucleic acid
A:Residues: 1-284 <HAR>
A:Cross-references: ENBL:X61273; NID:G64553; PIDN:CAA43577.1; PID:g64554
C:Superfamily: tropomyosin

Query Match 10.4%; Score 109; DB 2; Length 284;
Best Local Similarity 23.7%; Pred. No. 2.6;
Matches 49; Conservative 37; Mismatches 81; Indels 40; Gaps 8;

QY      1 MSKKKGLSAE---EKRTFMELFSETKDVFQ-----LKDLEKIAPK-EKGITAMSV 47

```

Db	73	LSDKKATDAEGDVASLNRRIQLVVEELDRAQERLSTALQLEEAEEKAADESERGMKV	132
Qy	48	KEVLQSLVDDGMYDCERIGTSNYWAPPSKALHARXHKLEVLESQLESGSQXHASLQXSI	107
Db	133	R-----ALKOEERKWELOEI-----QLKEAKHIAEEADRYEEVARKLVIIEGDL	176
Qy	108	EKAKIGRCETEER--TRLAKELSSLRDQREQLKAEVEKYKDCDPPQVVEEIRQANKVAK	165
Db	177	ERAE-ERAELESSEKCAELEBEELKXTVNNLKSLEAQAEKYSQKEDKYEEIEIKVLT	235
Qy	166	ANRTDNIFAISWAKRKGFEENKID	192
Db	236	ETR-----ABFAERTVAKLEKSID	254

Search completed: September 27, 2004, 08:39:29
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 08:36:32 ; Search time 12 Seconds
(without alignments)
889.532 Million cell updates/sec

Title: us-10-087-190-3
Perfect score: 1047
Sequence: 1 MSKKGLSABEKRNMETP.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	349	33.3	210	1	YA53_SCHPO
2	168.5	16.1	174	1	YGTS_YEAST
3	135.5	12.9	1281	1	DYNA_MOUSE
4	134.5	12.8	1278	1	DYNA_HUMAN
5	131.5	12.6	1280	1	DYNA_RAT
6	130.5	12.5	1224	1	DYNA_CHICK
7	121.5	11.6	1357	1	KTNI_HUMAN
8	121.5	11.6	1384	1	KTNI_CHICK
9	120	11.5	882	1	RA50_PRRFU
10	117	11.2	976	1	SCPI_HUMAN
11	116.5	11.1	2442	1	CEP2_HUMAN
12	115.5	11.0	880	1	RA50_PRRAB
13	114	10.9	284	1	TPM1_PIG
14	113	10.8	199	1	TPM1_YEAST
15	112	10.7	284	1	TPM3_HUMAN
16	112	10.7	1938	1	MYH4_RABIT
17	111	10.6	1937	1	MYH8_HUMAN
18	110.5	10.6	879	1	RA50_PRRHO
19	110.5	10.6	1433	1	REST_CHICK
20	110	10.5	284	1	TPM1_BRARE
21	110	10.5	284	1	TPM1_RANTE
22	110	10.5	284	1	TPM1_RANTE
23	109.5	10.5	1391	1	MST2_DROHY
24	109	10.4	284	1	TPM1_XENIA
25	109	10.4	1084	1	MYSS_METUA
26	107.5	10.3	1169	1	SMC_METUA
27	107.5	10.3	1197	1	SMC2_HUMAN
28	107	10.2	284	1	TPM1_HUMAN
29	107	10.2	284	1	TPM1_MOUSE
30	107	10.2	284	1	TPM1_RABIT
31	107	10.2	284	1	TPM1_RAT
32	107	10.2	1941	1	MYH2_HUMAN
33	106	10.1	1330	1	KTNI_VULVU

RESULT 1

34	106	10.1	2116	1	MYS2_DICDI
35	105.5	10.1	895	1	RA50_THEVO
36	104	9.9	558	1	YH32_AQUAE
37	104	9.9	1005	1	RA50_METUA
38	104	9.9	1939	1	MYH1_HUMAN
39	104	9.9	1940	1	MYH3_CHICK
40	103	9.8	473	1	YVCE_BACSU
41	103	9.8	886	1	RA50_ARCFU
42	103	9.8	1509	1	MYSN_ACACA
43	103	9.8	1935	1	MYSS_CYPCA
44	102.5	9.8	697	1	MFP1_LYCES
45	102	9.7	284	1	TPM1_CHICK
46	102	9.7	284	1	TPM1_COTJA
47	102	9.7	372	1	M22_STRPY
48	102	9.7	476	1	VSG4_TRYER
49	102	9.7	1031	1	KINH_STRPU
50	102	9.7	1940	1	MYH3_HUMAN
51	101.5	9.7	1976	1	MYHA_RAT
52	101.5	9.7	2663	1	CENE_HUMAN
53	101	9.6	804	1	MEAF_HUMAN
54	101	9.6	1225	1	KF4A_CHICK
55	101	9.6	1756	1	PEPL_HUMAN
56	100.5	9.6	1001	1	RFRG_MOUSE
57	100	9.6	1539	1	Y373_HUMAN
58	100	9.6	1549	1	SMC4_CAEBL
59	100	9.6	1938	1	MYSS_CHICK
60	99	9.5	284	1	TPM_PRRAM
61	99	9.5	539	1	M24_STRPY
62	99	9.5	978	1	RA50_AQUAE
63	99	9.5	1053	1	EM12_HUMAN
64	99	9.5	1130	1	Y117_CAEBL
65	99	9.5	1164	1	BAG_STRAG
66	98.5	9.4	1755	1	PEPL_MOUSE
67	98.5	9.4	1959	1	MYH9_CHICK
68	98	9.4	284	1	TPM2_CHICK
69	98	9.4	480	1	SIFA_METUA
70	98	9.4	1074	1	EM12_MOUSE
71	98	9.4	1453	1	Y373_BOVIN
72	98	9.4	2035	1	EVPL_MOUSE
73	98	9.4	3210	1	CENF_HUMAN
74	97.5	9.3	284	1	TPM2_HUMAN
75	97.5	9.3	284	1	TPM2_MOUSE
76	97.5	9.3	284	1	TPM2_RABIT
77	97.5	9.3	284	1	TPM2_RAT
78	97.5	9.3	539	1	MYS3_HYDAT
79	97.5	9.3	1085	1	YAF4_SCHPO
80	97.5	9.3	1286	1	SMC4_MOUSE
81	97.5	9.3	1976	1	MYHA_BOVIN
82	97.5	9.3	1976	1	MYHA_HUMAN
83	97.5	9.3	2017	1	MYSN_DROME
84	97	9.3	407	1	M21_STRPY
85	97	9.3	1044	1	YAF3_SCHPO
86	97	9.3	1790	1	USO1_YEAST
87	97	9.3	1934	1	MYH7_MESAU
88	97	9.3	2058	1	MY10_HUMAN
89	96.5	9.2	724	1	HMNR_HUMAN
90	96.5	9.2	964	1	YOY1_CAEBL
91	96.5	9.2	1191	1	CING_MOUSE
92	96.5	9.2	1243	1	SMC4_MICAR
93	96.5	9.2	1938	1	MYHD_HUMAN
94	96.5	9.2	1939	1	MYH4_HUMAN
95	96.5	9.2	2230	1	GOA4_HUMAN
96	96.5	9.2	4687	1	PLEI_RAT
97	96	9.2	248	1	TPM4_RAT
98	96	9.2	396	1	TRT_DROME
99	96	9.2	887	1	YLX8_CAEBL
100	96	9.2	1818	1	HWM2_MYCPN

ALIGNMENTS

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

EMBL; AF064205; AAD55811.1; -.
EMBL; AF064203; AAD55811.1; JOINED.
EMBL; AF064204; AAD55811.1; JOINED.
EMBL; AF064205; AAD55812.1; -.
EMBL; AF064206; AAD55812.1; -.
EMBL; AF064204; AAD55812.1; JOINED.
EMBL; X98801; CAA67333.1; -.
EMBL; AF086947; AAD03694.1; -.
EMBL; AF086927; AAD03694.1; JOINED.
EMBL; AF086928; AAD03694.1; JOINED.
EMBL; AF086929; AAD03694.1; JOINED.
EMBL; AF086930; AAD03694.1; JOINED.
EMBL; AF086931; AAD03694.1; JOINED.
EMBL; AF086932; AAD03694.1; JOINED.
EMBL; AF086933; AAD03694.1; JOINED.
EMBL; AF086934; AAD03694.1; JOINED.
EMBL; AF086935; AAD03694.1; JOINED.
EMBL; AF086936; AAD03694.1; JOINED.
EMBL; AF086937; AAD03694.1; JOINED.
EMBL; AF086938; AAD03694.1; JOINED.
EMBL; AF086939; AAD03694.1; JOINED.
EMBL; AF086940; AAD03694.1; JOINED.
EMBL; AF086941; AAD03694.1; JOINED.
EMBL; AF086942; AAD03694.1; JOINED.
EMBL; AF086943; AAD03694.1; JOINED.
EMBL; AF086944; AAD03694.1; JOINED.
EMBL; AF086945; AAD03694.1; JOINED.
EMBL; AF086946; AAD03694.1; JOINED.
GeneW; HGNC:2711; DCTN1.
MIM; 601143; -.
MIM; 607641; -.
GO; GO:0005737; C:cytoplasm; TAS.
InterPro; IPR000938; CAP-Gly.
Pfam; PF01302; CAP Gly; 1.
PROSITE; PS00845; CAP GLY 1; 1.
PROSITE; PS02045; CAP GLY 2; 1.
Motor protein; Microtubule; Dynein; Coiled coil; Cytoskeleton;
Alternative splicing; Phosphorylation; Polymorphism.
CAP-Gly.
DOMAIN 48 90
SER-RICH.
DOMAIN 164 191
COILED COIL (POTENTIAL).
DOMAIN 213 547
COILED COIL (POTENTIAL).
DOMAIN 943 1049
COILED COIL (POTENTIAL).
DOMAIN 1182 1211
COILED COIL (POTENTIAL).
VARSPLIC 1 138
MQSKHVVSPPTSGSRMSAASAPLFGVSRVEVIGKGR
GTVAYGVALFQATGWGVLDDEAKGNDGTVGQRKFTCD
GTVGVFVRSQLOVFEDGADVTSPETPPDSSASKVLKRGTD
TKTSKLRLKPKK -> MMRQ (in isoform
p135).
/FTid=VSP 000760.
G -> S (in PLMND).
/FTid=VAR_015850.
A -> P.
/FTid=VAR_001373.
S -> N (IN REF. 2 AND 3).
MISSING (IN REF. 2 AND 3).
D -> V (IN REF. 2 AND 3).
MW; 6DCEASE57856E4BC CRC64;
SEQUENCE 1278 AA; 4141694

```

QY	62	CERI	-----GTSNYWAFPSKALHAR-KHKLEVLSEQLSGSQKHASLOKSIEKA	110
DB	339	LEILKAEIEKSGDGAASGYQLKQEBEQNARKDALVNRDLSUSSEKQBHVKLQKLMEK-	397	
QY	111	KICRCET--EERTRIAKELSSLRDQEQKAEVEKYKQCDPQVVEIRQANKVAKAEANR	168	
DB	398	KNQELVVRQQRERLQEBLSQAESTIDELKEQVDAALGAE-ENVEMLTDRLNLLEEKVRE	456	
QY	169	WTDNTPFAKSWAKRFGFEENKIDRTFGIPEDFD	202	
DB	457	LSEETVGDLEANNMDELQENARETELELREQLD	490	
RESULT 5				
DYNA_RAT				
ID	DYNA_RAT	STANDARD;	PRT; 1280 AA.	
AC	P28023;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Dynactin 1 (150 kDa dynein-associated polypeptide) (DP-150) (DAP-150)			
DE	(p150-glued).			
GN	DCTN1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RC	MEDLINE=91260877; PubMed=1828935;			
RX	Holzbur E.L.F., Hammarback J.A., Paschal B.M., Kravit N.G.,			
RA	Pfister K.K., Vallee R.B.,			
RA	"Homology of a 150K cytoplasmic dynein-associated polypeptide with			
RT	the Drosophila gene Glued.";			
RL	Nature 351:579-580 (1991).			
RN	[2]			
RP	REVISIONS.			
RP	Holzbur E.L.F., Hammarback J.A., Paschal B.M., Kravit N.G.,			
RA	Pfister K.K., Vallee R.B.,			
RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: Required for the cytoplasmic dynein-driven retrograde			
CC	movement of vesicles and organelles along microtubules. Dynein-			
CC	dynein interaction is a key component of the mechanism of axonal			
CC	transport of vesicles and organelles.			
CC	!- SUBUNIT: Large macromolecular complex of at least 10 components;			
CC	p150 (glued) binds directly to microtubules and to cytoplasmic			
CC	dynein.			
CC	!- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	!- SIMILARITY: Belongs to the dyneactin 150 kDa subunit family.			
CC	!- SIMILARITY: Contains 1 CAP-Gly domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X62160; CAA44091.1; -			
DR	PIR; S16129; S16129.			
DR	InterPro; IPR000938; CAP-Gly.			
DR	Pfam; PF01302; CAP_Gly; 1.			
DR	PROSITE; PS00845; CAP_Gly_1; 1.			
DR	PROSITE; PS50245; CAP_Gly_2; 1.			
KW	Motor protein; Microtubule; Dynein; Coiled coil; Cytoskeleton.			
FT	DOMAIN 48 90			
FT	DOMAIN 157 184			
FT	DOMAIN 214 513			
FT	DOMAIN 942 1048			
FT	DOMAIN 1184 1213			
FT	SEQUENCE 1280 AA; 141929 MW; C9348CF129FAFF5C CRC64;			
SO	Motor protein; Microtubule; Dynein; Coiled coil; Cytoskeleton.			

DR	InterPro; IPR000938; CAP-Gly.	
DR	Pfam; PF01302; CAP_GLY; 1.	
DR	PROSITE; PS00845; CAP_GLY_1; 1.	
DR	PROSITE; PS0245; CAP_GLY_2; 1.	
KW	Motor protein; Microtubule; Dynein; Coiled coil; Cytoskeleton;	
KW	Alternative splicing.	
FT	DOMAIN 49 91	CAP-GLY
FT	DOMAIN 205 540	COILED COIL (POTENTIAL).
FT	DOMAIN 936 1042	COILED COIL (POTENTIAL).
FT	DOMAIN 1081 1117	COILED COIL (POTENTIAL).
SQ	SEQUENCE 1224 AA; 135562 MW; 03B7FFE68E7C01D7 CRC64;	
Query Match	12.5%; Score 130.5; DB 1; Length 1224;	
Best Local Similarity	23.6%; Pred. No. 0.37; 93; Indels 37; Gaps 7;	
Matches	53; Conservative 42; Mismatches 37; Gaps 7;	
QY	2 SKKKGSAEKRTMKEIFSETKVFQKDKLEKIAPKEGITAM-----SVKEVLQSLVDD 57	
DB	272 AKKAKDALEAKERYVEEMADTADATAMATLDKEMAEPAEASLQQEYDLSLKEKVEYLTMD 333	
QY	58 GMV---DCERIGTSNYWAFPSKALHARKHKLKLE---VLESQLEGG-SQKHASLQKSTKA 110	
DB	332 LEIKHIEIEKSGDGAASSVQVQLSEQNARLKEALVRMDLSASEKQEHVKLQKQHEKK 391	
QY	111 KIGRCETEETRLAKLSSLRDQRELKAEVEK-YKDCD-----PQVVEBIRQ 157	
DB	392 N-----TELESRLQQREKLQEEVQKAEKTVDELKCEQVDAALGAAEVETLITE 438	
QY	158 ANKVAKEAANRWTDNIFAIKSWAKRKGPEENKIDRTFGIPEDFD 202	
DB	439 RNLDLEEKVRELRETVDGLDLEAMNEMNDELQENARETELELRQOLD 483	
RESULT 7		
KTNL	HUMAN	
ID	KTNL_HUMAN STANDARD; PRT; 1357 AA.	
AC	Q86UP2; Q13999; Q14707; Q15387; Q86W57;	
DT	10-OCT-2003 (Rel. 42, Created)	
DT	10-OCT-2003 (Rel. 42, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Kinectin (Kinesin receptor) (CG-1 antigen).	
GN	KTNL OR CGI OR KIA0004.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A. (ISOFORM 1).	
RC	TISSUE=Lymphoid;	
RX	MEDLINE=95306853; PubMed=7787243;	
RA	Fueterer A.; Kruppa G.; Kraemer B.; Lemke H.; Kroenke M.;	
RT	"Molecular cloning and characterization of human kinectin.";	
RL	Mol. Biol. Cell 6:161-170(1995).	
RN	[2]	
RP	SEQUENCE FROM N.A. (ISOFORM 2).	
RC	TISSUE=Periphereal blood lymphocytes;	
RX	MEDLINE=94314220; PubMed=8039706;	
RA	Print C.G.; Leung B.; Harrison J.E.B.; Watson J.D.; Krissansen G.W.;	
RT	"Cloning of a gene encoding a human leukocyte protein characterised by	
RL	extensive heptad repeats.";	
RL	Gene 144:221-228(1994).	
RN	[3]	
RP	SEQUENCE FROM N.A. (ISOFORM 1).	
RA	Wang H.-C.; Chen W.-F.; Su Y.-R.;	
RT	"Identification of a variant of Homo sapiens kinectin mRNA.";	
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.	
RN	[4]	
RP	SEQUENCE FROM N.A. (ISOFORM 2).	
RC	TISSUE=Bone marrow;	
RX	MEDLINE=96051387; PubMed=7584026;	
RA	Nomura N.; Miyajima N.; Suzuki T.; Tanaka A.; Kawabayashi Y.;	
RA	Sato S.; Nagase T.; Seki N.; Ishikawa K.-I.; Tabata S.;	
RT	"Prediction of the coding sequences of unidentified human genes. I.	

The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by RT analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1.";
RNA Res. 1:27-35(1994).
[5]
RN SEQUENCE FROM N.A.
RX PubMed=42508121;
RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C., Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A., Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S., Sun H., Du H., Pepin K., Arthiguenave F., Robert C., Cruaud C., Brues T., Jaillon O., Friedlander L., Sanson G., Brothier P., Cure S., Segreus B., Aniere F., Smalin S., Crespeau H., Abbasi N., Alach N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman C., Gouyvenoux M., James R., Madan A., Malrey-Estrada B., Mangenot S., Martins N., Menard M., Ootas S., Ratcliffe A., Shafer T., Trask B., Vacherie B., Bellemere C., Belser C., Besnard-Gonnet M., Bartol-Mavel D., Boutard M., Brizet-Silla S., Combette S., Dufosse-Laurent V., Perron C., Lechaplais C., Louesse C., Muselet D., Magdelenat G., Pateau E., Petit E., Sirvain-Trukiewicz P., Trybou A., Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M., Dumont C., Guerin T., Haifry S., Hammadi R., Muanga J., Pellouin V., Robert D., Wunderle E., Gauguet G., Roy A., Sainte-Marthe L., Verdier J., Verdier-Discala C., Hillier L.W., Fulton L., McPherson J., Matsuda F., Wilson R., Scarpelli C., Gyapay G., Wincker P., Sautin W., Quetier P., Waterston R., Hood L., Weissbach J.,
RT "The DNA sequence and analysis of human chromosome 14.";
RL Nature 421:601-607(2003).
[6]
RN SEQUENCE OF 1-870 FROM N.A. (ISOFORMS 1/2).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Ratkin L., Mariani K., Farmer A.A., Rubin G.M., Hong L., Datchenko L., Marusina K., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallalón D.K., Murthy D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield A.S.N., Kzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
RN SEQUENCE OF 191-195; 395-405; 545-554; 747-761 AND 811-823.
RX MEDLINE=21969647; PubMed=11973345;
RA Tran H., Pankov R., Tran S.D., Hampton B., Burgess W.H., Yamada K.M.;
RT "Integrin clustering induces kinectin accumulation.";
RL J. Cell Sci. 115:2031-2040(2002).
[8]
RN CHROMOSOMAL LOCATION.
RX MEDLINE=96163023; PubMed=8575822;
RA Print C.G., Morris C.M., Spurr N.K., Rooke L., Kriessens G.W.;
RT "The CG-1 gene, a member of the kinectin and ES/130 family, maps to human chromosome band 14q22.";
RL Immunogenetics 43:227-229(1996).
CC -1- FUNCTION: Receptor for kinesin thus involved in kinesin-driven vesicle motility. Accumulates in integrin-based adhesion complexes (IAC) upon integrin aggregation by fibronectin.
CC -1- SUBUNIT: Parallel homodimers formed between the membrane-bound and the cytosolic form, and also between 2 cytosolic forms (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Vesicle membrane protein anchored to the endoplasmic reticulum.

CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q86UP2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q86UP2-2; Sequence=VSP_007981. VSP_007982;
CC TISSUE SPECIFICITY: High levels in peripheral blood lymphocytes, testis and ovary, lower levels in spleen, thymus, prostate, small intestine and colon.
CC -1- SIMILARITY: Belongs to the kinectin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: Z22551; CAA80271.1; -;
CC EMBL: L25616; AAB65853.1; -;
CC EMBL: AY264265; AAP20418.1; -;
CC EMBL: D13629; BAA02794.1; -;
CC EMBL: ALI38459; -; NOT ANNOTATED CDS.
CC EMBL: BC050855; AAH50355.1; ALT_TERM.
CC PIR: S32763; S32763.
CC PIR: I53799; I53799.
CC Genew; HGNC:6467; KTN1.
CC MIN; 600381; -;
CC GO: GO:0005789; C:endoplasmic reticulum membrane; TAS.
CC GO: GO:0005887; C:integral to plasma membrane; TAS.
CC GO: GO:0005624; C:membrane fraction; TAS.
CC GO: GO:0006859; P:nonselective vesicle transport; TAS.
CC GO: GO:0007018; P:microtubule-based movement; ISS.
CC InterPro: IPR02017; Spectrin.
CC Endoplasmic reticulum; Signal-anchor; Transmembrane; Coiled coil;
KW Alternative splicing; Polymorphism.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 29 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).
FT DOMAIN 30 1357 ENDOPLASMIC RETICULUM (POTENTIAL).
FT DOMAIN 330 1356 COILED COIL (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 772 772 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 904 904 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1088 1088 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1263 1263 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1329 1329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1031 1059 Missing (in isoform 2).
FT VARSPLIC 1232 1259 Missing (in isoform 2).
FT VARIAT 282 282 V -> M (in dbSNP:2274073).
FT CONFLICT 15 15 S -> P (IN REF. 4).
FT CONFLICT 210 210 MISSING (IN REF. 1).
FT CONFLICT 373 373 I -> M (IN REF. 1).
FT CONFLICT 939 939 E -> G (IN REF. 1).
SQ SEQUENCE 1357 AA; 156274 MW; 971FCDFAA8FC88E CRC64;
Query Match 11.6%; Score 121.5; DB 1; Length 1357;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 50; Conservative 88; Mismatches 51; Gaps 5;
OY 3 KKKGLSAAEERTRMEIFSETKDVFLQKLEKIAKPEKIGITA--MSVKEVLQSLVDGMV 60
DB 1026 RKKNDLRKNWEAMEALASTERKVLQDKVYKTSKERQQQVEAVELEAKVILKLPKVSU 1085
OY 61 DCEIGTSNTYWPAPPSVAH-----AKKKLVLESLSEGSOKHASLQKSIKAKIGRC 115

```

Db 1086 P-SNL5YGBWLAGFEKKAKCMAGTSGSEVVKLEHKLKAEADMTLLQLECEKYKSVLA 1144
QY 116 ETE-----ERELA-KELSSLRD 132
Db 1145 ETEGLQKLQSVQEBENKVKVDESHKTIYQMOSSFTSSQELERLSENKNDIENLR 1204
QY 133 QRELKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDN 172
Db 1205 EREHLEMELEKAEEMERSTVTEVRELKDLTLQKQLDSD 1244

RESULT 8
KTNI_CHICK
ID KTNI_CHICK STANDARD; PRT; 1364 AA.
AC Q90631;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kinetin.
GN KTN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2-12 AND 233-247.
RX MEDLINE=99041931; PubMed=9822636;
RA Kumar J., Erickson H.P., Sheetz M.P.;
RT "Ultrastructural and biochemical properties of the 120-kDa form of
chick Kinetin.";
RL J. Biol. Chem. 273:31738-31743(1998).
CC -I- FUNCTION: Receptor for kinesin thus involved in kinesin-driven
vesicle motility.
CC -I- SUBUNIT: Parallel homodimers formed between the membrane-bound and
the cytosolic form, and also between 2 cytosolic forms.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Vesicle membrane
protein anchored to the endoplasmic reticulum.
CC -I- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=1;
Comment=A number of isoforms are produced;
Name=1;
IsoId=Q90631-1; Sequence=Displayed;
CC -I- PTM: Both the membrane and cytoplasmic forms seem to be
myristoylated.
CC -I- MISCELLANEOUS: A cytoplasmic form lacking the first 232 amino
acids has been characterized.
CC -I- SIMILARITY: Belongs to the kinectin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U15617; AAA85818.1;
DR Endoplasmic reticulum; Signal-anchor; Transmembrane; Coiled coil;
KW Alternative splicing; Myristate.
DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
TRANSMEM 7 29 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 30 1364 ENDOPLASMIC RETICULUM (POTENTIAL).
FT DOMAIN 315 1085 COILED COIL (POTENTIAL).

```

```

FT DOMAIN 1116 1306 COILED COIL (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 704 704 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 775 775 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 976 976 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1088 1088 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1094 1094 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1364 AA; 155976 MW; 176BB1A19A80F00 CRC64;

Query Match 11.6%; Score 121.5; DB 1; Length 1364;
Best Local Similarity 22.3%; Pred. No. 1.6;
Matches 49; Conservative 34; Mismatches 86; Indels 51; Gaps 5;

QY 3 KKKGLSAREKEKTRMWEIFSETKDVFLKDLKLEKAPKEKGTAMSV--KEVLQSLVDDGMV 60
Db 1032 RKKNDLREKKNWKAMEALATEKLLQDKVNTAKEKQOHVEAEVETRELLQKLPFNVL 1091
QY 61 DCEIRIGTSNYWAFPSKALH-----ARKHKLEVLSEQLSEGSQKSHASLQKSIKAKIGRC 115
Db 1092 PA-NVSHSEWICGPEKMAKEYLRGASGSEDIKVMQKLEKEAEELHILLQLECEKYKSVLA 1150
QY 116 ETE-----ERTEL---AKELSSLRD 132
Db 1151 ETEGILQRLQSRVEEBSKMKIKVEESQKELKQNRSSVASLEHEVERLKEIKEVETLKK 1210
QY 133 QRELKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDN 172
Db 1211 EREHLESELEKAEIETERTYTYVSEVRELKDLTLQKQLDSD 1250

RESULT 9
RASO_PYRFU
ID RASO_PYRFU STANDARD; PRT; 882 AA.
AC P58301;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 42, Last annotation update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR PF1167.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=20485553; PubMed=11029422;
RA Hopfner K.-P., Karcher A., Shin D., Fairley C., Tainer J.A.,
Carney J.P.;
RA "Weill and Rad50 from Pyrococcus furiosus: cloning and biochemical
characterization reveal an evolutionarily conserved multiprotein
machine.";
RL J. Bacteriol. 182:6036-6041(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=20349838; PubMed=10892749;
RA Weis R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RN [3]
RP Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-149.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=20349838; PubMed=10892749;
RA Hopfner K.-P., Karcher A., Shin D.S., Craig L., Arthur L.M.,
Carney J.P., Tainer J.A.;
RT "Structural biology of Rad50 ATPase: ATP-driven conformational
control in DNA double-strand break repair and the ABC-Atrase
superfamily.";
RL Cell 101:789-800(2000).

```


FT CONFLICT 516 528 L7SHCNKLSLENK -> YFTLQASPPFN (IN REF. 2).

FT CONFLICT 549 549 N -> I (IN REF. 2).

FT CONFLICT 560 560 K -> T (IN REF. 2).

FT CONFLICT 805 805 E -> D (IN REF. 2).

FT CONFLICT 941 941 P -> S (IN REF. 2).

SQ SEQUENCE 976 AA; 114069 MW; 88A81D042AC2696B CRC64;

Query Match 11.2%; Score 117; DB 1; Length 976;
Best Local Similarity 23.9%; Pred. No. 2.1;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKGLGAEK-KRTRMEI-FSETDVF-----FQLKDLKIAKPKKGTAMSVKEVL 51
DB 413 OKKSSLEEMTLTNKKVEELKVLGKETLLYENKQFKIAELKG-TEQELIGLL 471

QY 52 QSL-----VDGMDVCERIGTSNYVAPPSKAL-----HARKHKLVELE-- 90
DB 472 QAREKEVHLELTQLTAITTSSEYQYSKEVKDLTELENEKLNKLTSTHCNKLSENKELT 531

QY 91 -----SOLSSGSKQHASLQSIKAKIGRCETETRLAKELSSLRDQREQ 136
DB 532 QETSDMTLBNKQEDINNKKQERMLKQIE--NLQETETQLRNELEYVREELKQKDE 589

QY 137 LKAEEVKYD-CD-----POVVEETROANKVAKK--AANRWTDNIPATK----- 177
DB 590 VKCKLDKSENCKLNKQVENKKNYIEELQENKALKKKGTAESKQLNVIELKVNKLELE 649

QY 178 -SWAKEKFG-----FEENKI 191
DB 650 LBSAKQKFGIEITDYQKEIEDKKI 673

RESULT 11

CEP2_HUMAN

ID CEP2_HUMAN STANDARD; PRT; 2442 AA.

AC Q9B73; O14812; O60588; Q9H450;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Centrosomal protein 2 (Centrosomal Nek2-associated protein 1) (C-NAP1)

DE (Centrosome protein 250) (Centrosome associated protein CEP250).

GN CEP2 OR CNAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

SEQUENCE FROM N.A. (ISOFORM 1), AND AUTOIMMUNE DISEASE.

RP CEP2_HUMAN

RC TISSUE=Cervical carcinoma;

RX MEDLINE=98165428; PubMed=9506584;

RA Mack G.J., Rees J., Sandblom O., Balczon R., Fritzler M.J., Rattner J.B.;

RA "Autoantibodies to a group of centrosomal proteins in human autoimmune sera reactive with the centrosome.";

RL Arthritis Rheum. 41:551-558(1998).

[2]

SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), INTERACTION WITH NEK2, AND SUBCELLULAR LOCATION DURING THE CELL CYCLE.

RP CEP2_HUMAN

RC TISSUE=Placenta;

RX MEDLINE=98311641; PubMed=9647649;

RA Fry A.M., Mayor T., Meraldi P., Stierhof Y.-D., Tanaka K., Nigg E.A.;

RA "C-Nap1, a novel centrosomal coiled-coil protein and candidate substrate of the cell cycle-regulated protein kinase Nek2.";

RL J. Cell Biol. 141:1563-1574(1998).

[3]

SEQUENCE FROM N.A.

RP CEP2_HUMAN

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosht K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Leivaalaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultston J.E., Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Placenta;

RX MEDLINE=2238257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Haile S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skaleka U., Smalls D.E., Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [5]

RP PHOSPHORYLATION, AND INTERACTION WITH NEK2 AND PPP1CA.

RX MEDLINE=21060765; PubMed=10880350;

RA Helps N.R., Luo X., Barker H.M., Cohen P.T.W.;

RT "NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase localized to centrosomes, is complexed to protein phosphatase 1.";

RL Biochem. J. 349:509-518(2000).

[6]

RP PHOSPHORYLATION DURING CELL CYCLE.

RX MEDLINE=22135747; PubMed=12140259;

RA Mayor T., Hacker U., Stierhof Y.-D., Nigg E.A.;

RT "The mechanism regulating the dissociation of the centrosomal protein C-Nap1 from mitotic spindle poles.";

RL J. Cell Sci. 115:3275-3284(2002).

CC -!- FUNCTION: Probably plays an important role in centrosome cohesion during interphase.

CC -!- SUBUNIT: Monomer and homodimer (Probable). Forms a complex in vitro with both Nek2 kinase and the PPP1C catalytic subunit of protein phosphatase 1 (Pp1).

CC -!- SUBCELLULAR LOCATION: Component of the core centrosome. In interphase cells, it specifically associates with the proximal ends of both mother and daughter centrioles. Associates with the centrosome in interphase cells. In mitotic cells, it dissociates from the mitotic spindle poles. At the end of cell division, it reaccumulates at centrosomes.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;


```
Db 257 RKGLEEKIVQIERSIEKKAKISELEIVKDIPLQKEKEVYKLGKPR----- 306
QY 54 LVDDGMVDCERIGTSNYWAPPSKAL-----HARKHLEVLSEQLSEGSQKXHASL----- 103
Db 307 --DEYESKLRLEKE-SKWESELKAIEBEVKEGKKERAESEIREKLSIEIKRLEELKPY 364
QY 104 -----OKSIEKAK-----IGRET--EERTRLAKELSSL-----RD 132
Db 365 VELEDAKQVQKQIERLARKLGLSPGEVLEKLESEKTERTEIEBAIKITTRIGOMEGE 424
QY 133 QROLKA--EVEYKDCDPVVEIRQANKVAKEAANRWTDNIFAISWAKKFGFEENK 190
Db 425 KNERMKAIEBLRKAQKPCVCGRELTEEHK--KELMERYTLEIKTIEBELKRTTE--EERK 481
QY 191 I 191
Db 482 L 482

RESULT 13
TPM1 PIG STANDARD; PRT; 284 AA.
ID TPM1 PIG AC P42639;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
GN TPM1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (9.0 ANGSTROMS).
RC TISSUE=Heart muscle;
RX MEDLINE=93021087; PubMed=1404362;
RA Whitby F.G., Kent H.M., Stewart F., Stewart M., Xie X., Hatch V.,
RA Cohen C., Phillips G.N. Jr.;
RT "Structure of tropomyosin at 9-A resolution.";
RL J. Mol. Biol. 227:441-452 (1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (7.0 ANGSTROMS).
RX MEDLINE=20114460; PubMed=10651038;
RA Whitby F.G., Phillips G.N. Jr.;
RT "Crystal structure of tropomyosin at 7-A resolution.";
RL Proteins 38:49-59 (2000).
CC -1- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
CC plays a central role, in association with the troponin complex, in
CC the calcium dependent regulation of vertebrate striated muscle
CC contraction. Smooth muscle contraction is regulated by interaction
CC with caldesmon. In nonmuscle cells is implicated in stabilizing
CC cytoskeleton actin filaments.
CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms may be produced;
CC Name=1;
CC IsoId=P42639-1; Sequence=Displayed;
CC -1- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC exhibits a prominent seven-residues periodicity.
CC -1- SIMILARITY: Belongs to the tropomyosin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X66274; CAA46986.1; -
CC PIR; S24972; S24972.
CC PDB; 1C1G; 11-FEB-00.
```

```
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00261; tropomyosin; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PROSITE; PS00326; TROPOMYOSIN; 1.
KW Muscle protein; Cytoskeleton; Actin-binding; Coiled coil; Acetylation;
KW Multigene family; Alternative splicing; 3D-structure.
FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
FT MOD_RES 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 284 AA; 32729 MW; 071AD459050C7F98 CRC64;

Query Match 10.9%; Score 114; DB 1; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.91;
Matches 52; Conservative 37; Mismatches 61; Indels 74; Gaps 11;

QY 1 MSKKKGLSAB---EKRTMMEIFSETKDVFLKLEKIAPEKIGITAMSVKVEQLSLVDD 57
Db 73 LAEKATDAEADVASLNRIQLFEELDRAQ-----ERLA-----TALQKLEAEKAADE 122
QY 58 ---GMVDCERIGTSNYWAPPSKALHAR---KHKLEVLSEQLSEGSQKH----- 100
Db 123 SERGM-----KVIESRAQDKDEKMEIQEIQLKEA--KHIAEDADRKYE 163
QY 101 -----ASLQKSIEKAKI--GRCEETERTLAKELSLRDRQRLKAEVKKYKDCD 148
Db 164 EVARKLVIIESDLERAEERAEELSEGKC-----AELEBELKTVTNLKSLEAQAKYSQKE 218
QY 149 PQVVEEIRQANKVAKEAANRWTDNIFAISWAKKFGFEENKID 192
Db 219 DKYEEIKVLSDKLKEAETR-----APEAERSVTKLEKSID 254

RESULT 14
TPM1 YEAST STANDARD; PRT; 199 AA.
ID TPM1 YEAST AC P17536;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tropomyosin 1.
GN TPM1 OR YNL079C OR N2332.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89195234; PubMed=2649250;
RA Liu H., Bretscher A.;
RT "Disruption of the single tropomyosin gene in yeast results in the
RT disappearance of actin cables from the cytoskeleton.";
RL Cell 57:233-242 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=96267764; PubMed=8701611;
RA Poehlmann R., Philippsen P.;
RT "Sequencing a cosmid clone of Saccharomyces cerevisiae chromosome XIV
RT reveals 12 new open reading frames (ORFs) and an ancient duplication
RT of six ORFs.";
RL Yeast 12:391-402 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Soler-Mira A., Saiz J.E., Ballesta J.P.G., Remacha M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Not known.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC exhibits a prominent seven-residues periodicity.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M25501; AAA35174.1; -;
 DR EMBL; X86470; CAA60179.1; -;
 DR EMBL; Z71355; CAA95953.1; -;
 DR PIR; A32183; A32183.
 DR GenOnline; 143085; -;
 DR SGD; S0005023; TPM1.
 DR GO; GO:0000141; C:actin cable (sensu Saccharomycetes); IDA.
 DR GO; GO:0000142; C:contractile ring (sensu Saccharomycetes); IDA.
 DR GO; GO:0007015; P:actin filament organization; IPI.
 DR GO; GO:0008154; P:actin polymerization and/or depolymerization; IGI.
 DR GO; GO:0007118; P:apical bud growth; IGI.
 DR GO; GO:0000283; P:establishment of cell polarity (sensu Sacch. . . IGI.
 DR GO; GO:0007119; P:isotopic bud growth; IGI.
 KW Cytoskeleton; Coiled coil; Repeat.
 FT DOMAIN 1 199 COILED COIL.
 SQ SEQUENCE 199 AA; 23541 MW; 845F3DB4387083BF CRC64;
 Query Match 10.8%; Score 113; DB 1; Length 199;
 Best Local Similarity 24.7%; Pred. No. 0.73;
 Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;
 Qy 7 LSAPKTRMWEIFSETKDYVQLKLEKIAPKKGITAMSVKVLQSLVDD-----GMV 60
 Db 13 LEABSQKVEYELKKNKDLQ-ENVE---KENQIKSLTVKN--QQLEDEIEKLEAGLS 65
 Qy 61 DCRIGSNVYWPSPKALHARKHL-----EVLESQISEG-----SOK 99
 Db 66 DSKTQEDNVKENQIKSLTVKQHLEBIEKLEAEKLSQSESHHLSQNNDFSK 125
 Qy 100 HASLQKSIKAKIGRCETERTR-----LAKELSSLDQREQLKAEVE-----KYDC 147
 Db 126 NQLEEDLESSTKLTETKLEKRESLDKADQLERRVAALKEEQEENKNEELTVKYEDA 185
 Qy 148 DPQVVE 153
 Db 186 KKELDE 191

RESULT 15

TPM3 HUMAN
 ID -TPM3 HUMAN STANDARD; PRT; 284 AA.
 AC P06753; F12324; Q969Q2; Q9NQHS;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tropomyosin alpha 3 chain (Tropomyosin 3) (Tropomyosin gamma).
 GN TPM3.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=86311274; PubMed=3018581;
 RA Reinach F.C.; McLeod A.R.;
 RA "Tissue-specific expression of the human tropomyosin gene involved in
 RL the generation of the trk oncogene.";
 RL Nature 322:648-650(1986).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=87066720; PubMed=3024106;
 RA McLeod A.R.; Houliker C.; Talbot K.;
 RA "The mRNA and RNA-copy pseudogenes encoding TM30nm, a human
 RL cytoskeletal tropomyosin.";
 RL Nucleic Acids Res. 14:8413-8426(1986).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=88332987; PubMed=3418707;
 RA Clayton L., Reinach F.C., Chumbley G.M., MacLeod A.R.;

RT "Organization of the hTmnm gene. Implications for the evolution of
 RT muscle and non-muscle tropomyosins.";
 RL J. Mol. Biol. 201:507-515(1988).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Colon cancer;
 RA Lin J.-C., Lin J.L.-C., Geng X., Das K.M.;
 RT "Identification and characterization of a novel tropomyosin isoform
 RT from a colon cancer cell line T84.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Bone, Kidney, and Skeletal muscle;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=86146854; PubMed=2869410;
 RA Martin-Zanca D., Hughes S.H., Barbacid M.;
 RT "A human oncogene formed by the fusion of truncated tropomyosin and
 RT protein tyrosine kinase sequences.";
 RL Nature 319:743-748(1986).
 RN [7]
 RP PARTIAL SEQUENCE.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=93162043; PubMed=1286667;
 RA Rasmussen H.H., van Damme J., Puyse M., Gesser B., Celis J.E.,
 RA Vandeckerckhove J.;
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel
 RT protein database of normal human epidermal keratinocytes.";
 RL Electrophoresis 13:960-969(1992).
 RN [8]
 RP VARIANT NEM1 ARG-8.
 RX MEDLINE=95218823; PubMed=7704029;
 RA Laing N.G., Wilton S.D., Akkari P.A., Dorosz S., Bourdy K., Haan E.;
 RA Kneebone C., Blumbergs P., White S., Watkins H., Love D.R., Haan E.;
 RT "A mutation in the alpha tropomyosin gene TPM3 associated with
 RT autosomal dominant nemaline myopathy.";
 RL Nat. Genet. 9:75-79(1995).
 RN [9]
 RP ERRATUM.
 RX MEDLINE=95392584; PubMed=7663526;
 RA Laing N.G., Wilton S.D., Akkari P.A., Dorosz S., Bourdy K., Haan E.;
 RA Kneebone C., Blumbergs P., White S., Watkins H., Love D.R., Haan E.;
 RL Nat. Genet. 10:249-249(1995).
 RN [10]
 RP CHARACTERIZATION OF VARIANT ARG-8.
 RX MEDLINE=20056158; PubMed=10587521;
 RA Michele D.E., Albayya F.P., Metzger J.M.;
 RT "A nemaline myopathy mutation in alpha-tropomyosin causes defective
 RT regulation of striated muscle force production.";
 RL J. Clin. Invest. 104:1575-1581(1999).
 CC -!- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
 CC Plays a central role, in association with the tropomyosin complex, in


```

QY      160 KYVKAANFTWDNIFAIKSWAKRFGFEENKIDRTGEGIPD 200
DB      230 DKLKEAETRAE---FAERVA-----KLEKTIDDLDD 258

RESULT 17
MYH4 RABBIT
ID MYH4_RABIT STANDARD; PRT; 1938 AA.
AC Q28641
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, skeletal muscle, juvenile.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
ON NCBI_TaxID=9986;
RX [1]
SEQUENCE FROM N.A.
STRAIN=New Zealand white; TISSUE=Skeletal muscle;
MAEDA K., Hosinova E., Roesch-Kleinkauf A., Schuster H., Gasperik J.,
Wittinghofer A.;
RA "Isolation, sequencing of myosin heavy chain cDNA from rabbit
sketal muscle and a novel cosynthesis of S-1 fragment with the
essential and regulatory light chains.";
RT Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RL
CC -!- FUNCTION: Muscle contraction.
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
and 2 regulatory light chain subunits (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
characteristic for alpha-helical coiled coils.
CC -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
meromyosin (LMW) and 1 heavy meromyosin (HMW). It can later be
split further into 2 globular subfragments (S1) and 1 rod-shaped
subfragment (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licensesh-sib.ch).
-----
DR EMBL; U32574; AAA74199.1; --
DR FIC; A52933; A52933.
DR HSP; FI3538; ZMY3.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; Myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ_2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Multigene family.
FT DOMAIN 1 783 MYOSIN HEAD-LIKE.
FT DOMAIN 784 813 IQ.
FT DOMAIN 842 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 658 680 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 760 774 ACTIN-BINDING (BY SIMILARITY).

```

```

FT MOD RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
FT MOD RES 130 130 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
FT MOD RES 756 756 METHYLATION (TRI-) (BY SIMILARITY).
SQ SEQUENCE 1938 AA; 223064 MW; D8A8A2ECSB182626 CRC64;

Query Match 10.7%; Score 112; DB 1; Length 1938;
Best Local Similarity 22.2%; Pred.No.9.2; Indels 42; Gaps 7;
Matches 51; Conservative 44; Mismatches 93;

QY 1 MSKKGLSAAEKTRMEIFSETKQVFLQKD--LEKIAPKEKGITAMS-----VKREV 50
Db 1364 MSKANGAEVQWRKYETDAIORTTEEELEAKKLAQRLQDAEHEVAVNAKASLEKTKQR 1423
QY 51 LQSLVDGMDVDCRIGTS-----NYWAFPSKALHARKKLEVLSEQLSEGSQKHA 101
Db 1424 LQNEVEDLMDIVERTNAACALDKQRF-----DKILAERKHKYETHAELEASQKER 1478
QY 102 SLQSKIEKAK-----IGRCET--ERTRLAKELSLRDO-----RQLKAEVEK 143
Db 1479 SLSTEVFKVKNAYEESLDQLETLKRENKNLQQEISDLTSLQIABGGKRIHELEKVKQVQEQ 1538
QY 144 YKCDPQVVEIRQANKVAKEANRTDNIFAIKSWAKRFGPEENKIDR 193
Db 1539 EKSELQAALEAEASLEHEEGKILRIQLLENOVKSIDRKIAEKDEIDQ 1588

RESULT 18
MYH8_HUMAN
ID MYH8_HUMAN STANDARD; PRT; 1937 AA.
AC P13535; Q14910;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).
GN MYH8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Skeletal muscle;
RC MEDLINE=90323631; PubMed=2373371;
RA Karsch-Mizrachi I., Feghali R., Shows T.B. Jr., Leinwand L.A.;
RT "Generation of a full-length human perinatal myosin heavy-chain-
RT encoding cDNA.";
RL Gene 89:289-294(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95324556; PubMed=7601129;
RA Jullian B.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RA Stedman H.H., Rubinstein N.A.;
RT "Characterization of a human perinatal myosin heavy-chain
RT transcript.";
RL Eur. J. Biochem. 230:1001-1006(1995).
RN [3]
RP SEQUENCE OF 502-1937 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RT myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]
RP SEQUENCE OF 860-1937 FROM N.A.
RX MEDLINE=89234168; PubMed=2715179;
RA Feghali R., Leinwand L.A.;
RT "Molecular genetic characterization of a developmentally regulated
RT human perinatal myosin heavy chain.";
RL J. Cell Biol. 108:1791-1797(1989).

```

```

RN [5]
RP SEQUENCE OF 1-46 FROM N.A.
RA Esser K., Tidhar A., Myszkowski M.;
RT "Isolation and characterization of the human perinatal MHC promoter.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Muscle contraction.
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC and 2 regulatory light chain subunits (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMW) and 1 heavy meromyosin (HMM). It can later be
CC split further into 2 globular subfragments (S1) and 1 rod-shaped
CC subfragment (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M36769; AAC17185.1; -
CC EMBL; Z38133; CAA86293.1; -
CC EMBL; X51592; CAA35941.1; -
CC EMBL; AF067143; AAC21557.1; -
CC PIR; I38055; I38055.
CC HSP; P13538; 2MYS.
CC Genes; HGNC:7578; MYH8.
CC MIM; 160741; -
CC GO; GO:0005859; C:muscle myosin; TAS.
CC GO; GO:0008307; F:structural constituent of muscle; TAS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; Myosin head.
CC InterPro; IPR004009; Myosin N.
CC InterPro; IPR002928; Myosin tail.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MSC; 1.
CC PROSITE; PS50096; IQ; 1.
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC ATP-binding; Methylation; Multigene family; Calmodulin-binding.
CC
CC -----
CC DOMAIN 1 780
CC MYOSIN HEAD-LIKE.
CC
CC FT DOMAIN 781 813
CC FT DOMAIN 842 1937
CC FT NP_BIND 181 188
CC FT DOMAIN 658 680
CC FT DOMAIN 760 774
CC FT MOD_RES 132 132
CC FT CONFLICT 15 15
CC FT CONFLICT 970 970
CC FT CONFLICT 1072 1072
CC FT CONFLICT 1247 1247
CC FT CONFLICT 1251 1252
CC FT CONFLICT 1261 1261
CC FT CONFLICT 1297 1297
CC FT CONFLICT 1377 1378
CC FT CONFLICT 1504 1505
CC FT CONFLICT 1847 1847
CC FT CONFLICT 1914 1914
CC SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;
SQ

```

Query Match 10.6%; Score 111; DB 1; Length 1937;
 Best Local Similarity 23.1%; Pred. No. 11;
 Matches 54; Conservative 42; Mismatches 88; Indels 50; Gaps 9;

QY 4 KKGLSABEKTRMEIEFSETKDVFPOLKLEKIAPK-----EKGITAMS-----V 47
 DB 1361 ORALSKANSEVAQWRTRYETDAIQTEELBEAKKLAQRLQEAEEHVEAVNAKCALEKT 1420
 QY 48 KEVLSQVLDGWDVCEBRTGS-----NYWAPPSKALHARKHKLVLVESQLSEGSQ 98
 DB 1421 KQRLQNEVEBLMLDVSNACALDKKQNF-----DKVLSWKQYETQAELEASQK 1475
 QY 99 KHASLQKSIKAK-----IGRCET--ERTRIAKELSLRDQ-----REOLKAE 140
 DB 1476 ESRSLSTELFKVKNVYESLDQLETLRRNKNLQEIISDLTEQIABGGKQIHELEXIKQ 1535
 QY 141 VEKYKCDPQVWEIRANKVAKEA--ANRWTDNIPAIKSWAKKFGFEENKIDR 193
 DB 1536 VEGEK-CEIQAALEEAFALEHEEGKILRIQLBQNVKSEVDRIKAEKDEIDQ 1588

RESULT 19
 RA50_PVRHO STANDARD; PRT; 879 AA.
 AC OS8687;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR PH0929.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Osuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RA "Complete sequence and gene organization of the genome of a hyper-
 PT thermophilic archaebacterium, Pyrococcus horikoshii OT3";
 RL DNA Res. 5:55-75(1998).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -!- SIMILARITY: Belongs to the SMC family. RAD50 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP000004; BAA30025.1; -.
 CC PIR; C71083; C71083.
 CC HAMAP; MF_00449; -; 1.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR003439; ABC_transporter.
 CC InterPro; IPR007517; Rad50_zn_hook.
 CC InterPro; IPR003405; SMC_C.
 CC InterPro; IPR003395; SMC_N.
 CC Pfam; PF04423; Rad50_zn_hook; 1.

DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF02463; SMC_N; 1.
 DR SMART; SM00382; AAA; 1.
 KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 FT NP_BIND 30 37 ATP (BY SIMILARITY)
 FT DOMAIN 141 744 COILED COIL (POTENTIAL).
 SQ SEQUENCE 879 AA; 103673 MW; FD4E30FD1B8CB29 CRC64;

Query Match 10.6%; Score 110.5; DB 1; Length 879;
 Best Local Similarity 22.1%; Pred. No. 5;
 Matches 48; Conservative 51; Mismatches 77; Indels 41; Gaps 9;

QY 4 KKGLSABEKTRMEIEFSETKDVFPOLKLEKIAPK-----EKGITAMSKEVLQSLVDDG- 58
 DB 569 KKELS--EIEDRLRLGFKTIDELSGRIEELKPHNYIEAKNAEKRLDILSLKDERE 626
 QY 59 -----MVDCEIRIGT--SNYWAPPSKALHARKH-----LVLESQSEGS 97
 DB 627 ELDKAFELAKIETDIEKVTSQLNELQKFDOKYEEKREKMKLSMEIKGLETKLELE 686
 QY 98 QKHASLQKSIKAKIGRCETEERTRLAKELSLRDQREQLKAEVKEKYKCDPQVWEIRQ 157
 DB 687 RRDEIKSTIEKLEEKERKESAKVELEKLNIAIKRIEELRGKIKEYKALIKE--EALNK 744
 QY 158 ANKYAKAANRWTDNIP---AIKSWAKKFGFEENKI 191
 DB 745 IGEIASEIFSEFTDGKYGIAIRA-----EDNKV 773

RESULT 20
 REST_CHICK STANDARD; PRT; 1433 AA.
 AC 042184; 042228; 057563; 057564;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Restin (Cytoplasmic linker protein-170) (CLIP-170).
 GN RSN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98137792; PubMed=9469933;
 RX Gripatic L., Volosky J.M., Keller T.C. III;
 RT "Cloning and expression of chicken CLIP-170 and restin isoforms";
 RL Gene 206:195-208(1998).
 RN [2]
 RP SEQUENCE OF 17-1139 FROM N.A. (ISOFORMS 3 AND 4).
 RC TISSUE=Pectoralis muscle;
 RX MEDLINE=99002898; PubMed=9784600;
 RX Gripatic L., Keller T.C. III;
 RT "Identification and expression of two novel CLIP-170/Restin isoforms
 RT expressed predominantly in muscle";
 RL Biochim. Biophys. Acta 1405:35-46(1998).
 CC -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=042184-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=042184-2; Sequence=VSP_000761;
 CC Name=3; Synonyms=CLIP-170(11); Sequence=VSP_000763;
 CC IsoId=042184-3; Sequence=VSP_000762; VSP_000763;
 CC Name=4; Synonyms=CLIP-170(11+35);
 CC IsoId=042184-4; Sequence=VSP_000764;
 CC -!- SIMILARITY: Contains 2 CAP-Gly domains.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AF014012; AAC60344.1; -
DR ENBL; AF020764; AAC60345.1; -
DR ENBL; AF045650; AAC03547.1; -
DR ENBL; AF045651; AAC03548.1; -
DR InterPro; IPR000938; CAP-GLY.
DR InterPro; IPR001878; Znf CCHC.
DR Pfam; PF01302; CAP GLY; 2
DR SMART; SM00343; Znf C2HC; 1.
DR PROSITE; PS00845; CAP_GLY_1; 2.
DR PROSITE; PS0245; CAP_GLY_2; 2.
KW Cytoskeleton; Microtubule; Coiled coil; Repeat; Alternative splicing.
FT DOMAIN 79 121 CAP-GLY 1.
FT DOMAIN 144 207 SER-RICH.
FT DOMAIN 235 277 CAP-GLY 2.
FT DOMAIN 305 332 SER-RICH.
FT DOMAIN 351 1353 COILED COIL (POTENTIAL).
FT DOMAIN 1414 1427 CCHC-BOX.
FT VARSPPLIC 458 492 Missing (in isoform 2).
FT VARSPPLIC 458 492 /FTId-VSP 000761.
FT VARSPPLIC 458 492 TOPKLEHARTKEQLLFEKTKADKLOLEDTR -> RK
FT VARSPPLIC 458 492 ROISEDPEN (in isoform 3).
FT VARSPPLIC 458 492 /FTId-VSP 000762.
FT VARSPPLIC 458 492 S -> GGSKVS (in isoform 3).
FT VARSPPLIC 458 492 /FTId-VSP 000763.
FT VARSPPLIC 458 492 T -> RKQISEDPENT (in isoform 4).
FT CONFLICT 309 309 K -> R (IN REF. 2; AAC03547).
FT CONFLICT 440 440 E -> V (IN REF. 2; AAC03548).
FT SEQUENCE 1433 AA; 161026 MW; 5631CE8683498E23 CRC64;
Query Match 10.6%; Score 110.5; DB 1; Length 1433;
Best Local Similarity 24.7%; Pred. No. 8.3;
Matches 49; Conservative 40; Mismatches 62; Indels 47; Gaps 9;
QY 7 LSBEKTRMWEIPESTKDFQKLEKTAPEKKGITAMSVK-----VLQSLVDGMV 60
DB 875 MSSELQKSLNLTWETK-----LKEEE---REQUTEAKVLENDIAELMKS----- 920
QY 61 DCERIGTSNYWAFPSKALHARKHKLHVELESQSGSQHASLQKSGIE----KAKIGRCE 116
DB 921 -----SGDSSAQLMKMDELRLKRLERLEIQLELTKANERAVQLQKNVEQTAKAESQOE 976
QY 117 T--BERTRLAKELSLRDQEQKAEVEKYKCDPQVVEIRQANKVAKAANRWTDNIF 174
DB 977 TLKTHQBELKQMDQLTDMKKOMETSONQYKDL-----QA-KYKETSEMITKHDA 1026
QY 175 AIKSWAKRKGFENKID 192
DB 1027 DIK-----GFKQLLD 1037
RESULT 21
TPM1 BRARE STANDARD; PRT; 284 AA.
AC P13104;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
GN TPM1 OR TPMA.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
```

```
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345529; PubMed=2789276;
RA Chara O., Dorit R.L., Gilbert W.;
RT "One-sided polymerase chain reaction: the amplification of cDNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5673-5677(1989).
CC -!- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
CC Plays a central role, in association with the troponin complex, in
CC the calcium dependent regulation of vertebrate striated muscle
CC contraction. Smooth muscle contraction is regulated by interaction
CC with caldesmon. In nonmuscle cells is implicated in stabilizing
CC cytoskeleton actin filaments.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC -!- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC exhibits a prominent seven-residues periodicity.
CC -!- SIMILARITY: Belongs to the tropomyosin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M24635; AAA50021.1; -
DR PIR; I51731; I51731.
DR ZFIN; ZDB-GENE-990415-269; tpma.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00261; Tropomyosin.1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PROSITE; PS00326; TROPOMYOSIN; 1.
KW Muscle protein; Cytoskeleton; Actin-binding; Coiled coil.
SQ SEQUENCE 284 AA; 32722 MW; 7B6403B873CEE6BD CRC64;
Query Match 10.5%; Score 110; DB 1; Length 284;
Best Local Similarity 23.3%; Pred. No. 1.6; Indels 52; Gaps 10;
Matches 51; Conservative 44; Mismatches 72;
QY 1 MSKKKGLSAE---EKRTMMEIFSETKDVQF-----LKDLEKIAPK-EKGITAMSV 47
DB 73 LAEKATDAEGDVASLNRRILQVEEELDRAQLERLATALQLEEAKAADSESRGMKVIE 132
QY 48 KEVQLSLVDGMVDCERIGTSNYWAFPSKALHARKHKLHVELESQSGSQ-----HASL 103
DB 133 R-----ALKDEEKMLQEI-----QLKEAKHAEADRYEVARKLVIVGEL 176
QY 104 QKSIEKAKI--GRCTERTRLAKELSLRDQEQKAEVEKYKCDPQVVEIRQANKV 161
DB 177 ERTERAELNGKC-----SELEELKVTNNMKSLEAQAEKYSAKEDKYEEIKVLTDK 231
QY 162 AKEAANRWTDNIFAIKSWAKRKGFENKIDRTFGIPED 200
DB 232 LKEAETRAE---FAERSVA-----KLEKTIDDLLED 258
RESULT 22
TPM1 RANTE STANDARD; PRT; 284 AA.
AC P13105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
GN TPM1.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=89345529; PubMed=2788276;
 RA Ohara O., Dorit R.L., Gilbert W.;
 RL "One-sided polymerase chain reaction: the amplification of cDNA.";
 CC Proc. Natl. Acad. Sci. U.S.A. 86:5673-5677(1989).
 CC -!- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
 CC Plays a central role, in association with the troponin complex, in
 CC the calcium dependent regulation of vertebrate striated muscle
 CC contraction. Smooth muscle contraction is regulated by interaction
 CC with caldesmon. In nonmuscle cells is implicated in stabilizing
 CC cytoskeleton actin filaments.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -!- DOMAIN: The molecule is in a coiled coil structure. The sequence
 CC exhibits a prominent seven-residues periodicity.
 CC -!- SIMILARITY: Belongs to the tropomyosin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M24634; AAA18096.1; -;
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00261; Tropomyosin; 1.
 DR PRINTS; PR00194; TROPOMYOSIN.
 DR PROSITE; PS00326; TROPOMYOSIN; 1.
 KW Muscle protein; Cytoskeleton; Actin-binding; Coiled coil.
 FT
 FT SEQUENCE 284 AA; 32664 MW; D03680D9FC3D255 CRC64;
 SQ
 Query Match 10.5%; Score 110; DB 1; Length 284;
 Best Local Similarity 23.2%; Pred. No. 1.6;
 Matches 48; Conservative 39; Mismatches 80; Indels 40; Gaps 8;

QY 1 MSKKKGSLAE---EKRTWMEIFSETKDVQFQ-----LKDLKIAPK-EKGITAMSV 47
 DB 73 LAEKATADADVASLNRRLQVLEELDRAQERLALQKLEAEKAADESGMKVLEN 132
 QY 48 KEVLQSLVDGMVDCERIGTSNYWAFPSKALHARKHKLVLSEQLSEGSQKHSLSQSI 107
 DB 133 R---ALKDEKTELQEI-----QLKEAKHIAEADRYEEVARKLVITIEGDL 176
 QY 108 EKAKIGRCETEER--TFLAKELSLRDQREQLKAEVKKDCDPQVVEIRQANKVAKEA 165
 DB 177 ERAE-ERAELESEKCALESELTNTNGLKSLAQAEKYSQKEDKYEEIKVLTDKLKEA 235
 QY 166 ANEWTNIFAISWAKRKFGEENKID 192
 DB 236 ETR-----AFAETVAKLEKSID 254

RESULT 23
 MST2 DROHY
 ID MST2 DROHY STANDARD; PRT; 1391 AA.
 AC Q08696;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Axoneme-associated protein mat101(2).
 GN MST101(2).
 OS Drosophila hydei (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7224;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=95045538; PubMed=7957199;
 RA Neesen J., Padmanabhan S., Buenemann H.;
 RT "Randomly arranged repeats of a novel highly charged 16-amino-acid
 RT motif representing the major component of the sperm-tail-specific
 RT axoneme-associated protein family Dhmst101 form extended

RT alpha-helical rods within the extremely elongated spermatozoa of
 RT Drosophila hydei.";
 RL Eur. J. Biochem. 225:1089-1095(1994).
 CC -!- FUNCTION: Possible structural role in the sperm tail.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Testis. Primary spermatocytes and early
 CC spermatids.
 CC -!- DOMAIN: The predominant structure is alpha-helical.
 CC -!- POLYMORPHISM: Length polymorphisms exist between different
 CC strains, most likely caused by length variations within the tandem
 CC repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X73481; CAA51876.1; -;
 DR PIR; S51364; S51364.
 DR FlyBase; FBgn0020733; Dhyd\mst101(2).
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005198; F:structural molecule activity; IEP.
 DR GO; GO:0007288; P:axoneme assembly; IEP.
 KW Sperm; Repeat; Multigene family; Polymorphism.
 FT DOMAIN 332 1268
 FT [KR]-K-X-C-X-X-X-A-K-X-X-K-X-X-X-X-E.
 FT SEQUENCE 1391 AA; 159000 MW; 1B2A368F30F48878 CRC64;
 SQ
 Query Match 10.5%; Score 109.5; DB 1; Length 1391;
 Best Local Similarity 25.8%; Pred. No. 9.3;
 Matches 56; Conservative 33; Mismatches 83; Indels 45; Gaps 10;

QY 1 MSKKKGSLAEKRTWMEIFSETKDVQFQLEKIAPEK-----GTAMSVKEVLQ 52
 DB 770 LAKKK--AAEKKKCK--EAAKKEAAEKKKCKTAKKKEAAEKKKCKTAKKKEAAE 825
 QY 53 SLVDGMVDCERIGTSNYWAFPSKALHARKHKLVLSEQLSEGS---QKHSLSQSI 109
 DB 826 K-----KKCKAAKKEAAEKKKCKTAKKKEAAEKKKCKTAKKKEAAEKKKCK 879
 QY 110 AKIGRCETEERTRIA-----KELSLRDQREQLK--AEVEKYKDCDPQVVEIRQA--N 159
 DB 880 AAKRKEAAEKKCAEAAKKEKLEAAKKEAAEKKEAAKKEVAERKKCE-ELAKKIKKAAEK 938
 QY 160 KVAKEAANRWTDNIFAISWAKRKFGEENKIDRTFG 196
 DB 939 KKCKKLAK-----KEKKAGEKNLKKKAG 962

RESULT 24
 TP01 XENLA
 ID TP01 XENLA STANDARD; PRT; 284 AA.
 AC Q01173; Q01174;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 CX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Embryo, and Oocyte;
 RX MEDLINE=92104164; PubMed=1840524;
 RA Hardy S., Fishman M., Osborne H.B., Thiebaud P.;
 RT "Characterization of muscle and non muscle Xenopus laevis tropomyosin
 RT mRNAs transcribed from the same gene. Developmental and
 RT tissue-specific expression.";


```

RL  Eur. J. Biochem. 202:431-440(1991).
CC  -|- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
CC  plays a central role, in association with the troponin complex, in
CC  the calcium dependent regulation of vertebrate striated muscle
CC  contraction. Smooth muscle contraction is regulated by interaction
CC  with caldesmon. In nonmuscle cells is implicated in stabilizing
CC  cytoskeleton actin filaments.
CC  -|- SUBUNIT: Heterodimer of an alpha and a beta chain.
CC  -|- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=1; Synonyms=Muscle;
CC  IsoId=Q01173-1; Sequence=Displayed;
CC  Name=2; Synonyms=Nonmuscle;
CC  IsoId=Q01173-2; Sequence=VSP_006612, VSP_006613, VSP_006614;
CC  -|- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC  exhibits a prominent seven-residues periodicity.
CC  -|- SIMILARITY: Belongs to the tropomyosin family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X61273; CAA43577.1; -.
DR  EMBL; X61272; CAA43576.1; -.
DR  PIR; S19690; S19690.
DR  PIR; S19691; S19691.
DR  InterPro; IPR000533; Tropomyosin.
DR  Pfam; PF00261; Tropomyosin.
DR  PRINTS; PR00194; TROPOMYOSIN.
DR  PROSITE; PS00336; TROPOMYOSIN; 1.
KW  Muscle protein; Cytoskeleton; Actin-binding; Coiled coil;
KW  Alternative splicing; Multigene family.
FT  VARSPPLIC 1 80
FT  MDAIKKMKMLDKENALDRQAEADKKGAEDKSKOLED
FT  ELVALQKLTGTEDELQKYSEALKQAEKLELSDKKATD
FT  -> MAGITSLKAVKRRKIKCLQQAQDAEBAEKLOERDME
FT  RKUREA (in isoform 2).
FT  /FTIGVSP_006612.
FT  VARSPPLIC 189 211
FT  KCALSEELTKVTNNLKSLEAQA -> HYRQLEDQQRINDQ
FT  TKTLHASE (in isoform 2).
FT  /FTIGVSP_006613.
FT  VARSPPLIC 258 284
FT  DELYAOKLYKAISSELDHALNDMTSI -> EKVAHAKKEN
FT  LNMHQLDQTLLELNM (in isoform 2).
FT  /FTIGVSP_006614.
SQ  SEQUENCE 284 AA; 32650 MW; B7F6844D9900FB69 CRC64;

Query Match 10.4%; Score 109; DB 1; Length 284;
Best Local Similarity 23.7%; Pred. No. 1.9;
Matches 49; Conservative 37; Mismatches 81; Indels 40; Gaps 8;

QY 1 MSKKKGLSAR---EKTRMEIFSEYQVDFQ-----LKDLEKIAPK-BKGTAMSV 47
Db 73 LSDKKATADGVDVASLNRIQLVEELDRQERLSTALQKLEAEKADESERGMKVSN 132
QY 48 KEVLQSLVDGMDVCERIGTSNYWAPPKALHARKHKLVELESQSGQSHASLOKSI 107
Db 133 R-----ALKDEEKMELEQI-----QLKEAGHAEADRKVEEVARKLVIEGDL 176
QY 108 EKAKIGRCETFER--TRLAKELSSLRDREQLKAEVYKYKDCDPQVVEIRQANKVAKEA 165
Db 177 ERAE-ERAELESKCAELSEELXTVTNNLKSLEAQAQEKYSQEDKYVEEIKVLDTKLKEA 235
QY 166 ANRWTDNIFAIKSWAKKGTGFENKID 192
Db 236 ETR-----ABFAERTVAKLEKSID 254

RESULT 25
MYSS RABIT
ID MYSS_RABIT STANDARD; PRT; 1084 AA.

```

```

AC  P02562;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Myosin heavy chain, skeletal muscle (Fragments).
OS  Oryctolagus cuniculus (Rabbit).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX  NCBI_TaxID=9986;
RN  [1]
RP  SEQUENCE OF 1-258.
RA  Capony J.-P., Eizinga M.;
RT  "The amino acid sequence of A 34,000 dalton fragment from S-2 of
RT  myosin.";
RL  Biophys. J. 33:148A-148A(1981).
RN  [2]
RP  SEQUENCE OF 259-428.
RX  MEDLINE=85131142; PubMed=3972832;
RA  Lu R.C., Wong A.;
RT  "The amino acid sequence and stability predictions of the hinge
RT  region in myosin subfragment 2.";
RL  J. Biol. Chem. 260:3456-3461(1985).
RN  [3]
RP  SEQUENCE OF 409-1084 FROM N.A.
RX  MEDLINE=87304245; PubMed=3305014;
RA  Maeda K., Szakiel G., Wittinghofer A.;
RT  "Characterization of cDNA coding for the complete light meromyosin
RT  portion of a rabbit fast skeletal muscle myosin heavy chain.";
RL  Eur. J. Biochem. 167:97-102(1987).
CC  -|- FUNCTION: Muscle contraction.
CC  -|- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC  heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC  and 2 regulatory light chain subunits of the myofibrils.
CC  -|- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC  -|- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC  cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC  characteristic for alpha-helical coiled coils.
CC  -|- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC  meromyosin (LMW) and 1 heavy meromyosin (HMM). It can later be
CC  split further into 2 globular subfragments (S1) and 1 rod-shaped
CC  subfragment (S2).
CC  -|- SIMILARITY: Contains 1 myosin-like globular head domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X05958; CAA29391.1; -.
DR  PIR; A02985; A02985.
DR  PIR; A05280; A05280.
DR  PIR; S00084; S00084.
KW  Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW  ATP-binding; Multigene family.
FT  NON TER 1 1
FT  DOMAIN <1 >258 ALPHA-HELICAL TAILPIECE (S2).
FT  NON CONS 258 259
FT  DOMAIN <259 1084 RODLIKE TAIL (S2 AND LMW DOMAINS).
FT  VARIANT 455 1084 COILED COIL (POTENTIAL).
FT  VARIANT 405 405 L -> V.
FT  VARIANT 408 408 V -> L.
FT  VARIANT 421 421 E -> D.
FT  VARIANT 423 423 S -> G.
FT  VARIANT 426 426 K -> R.
SQ  SEQUENCE 1084 AA; 125488 MW; 229CFD69A6E1F7F0 CRC64;

Query Match 10.4%; Score 109; DB 1; Length 1084;
Best Local Similarity 24.7%; Pred. No. 7.7;
Matches 53; Conservative 35; Mismatches 79; Indels 48; Gaps 9;

```

QY 22 ETQVDFQKDLKLEKIAK-----RKGITAMS-----VKVQLSLVDGMVDCERI 65
DB 525 ETDAIORTLEBAKAKLQRLQDAEHEVAVNSKCSLEKTKQRLQNEAEDLMIDVERS 584
QY 66 GTS-----NYWAFPSKALHAKHKLVLVLSQSGSKHASLQSIKAK----- 111
DB 585 NATCARMDKQKQNF-----DKVLAEWKHYEETQAELEASQKESRSLSTFVKVKNAYEE 639
QY 112 -IGRCET--EERTRLAKESSLRDQ--RQOLK--AEVEKQCDQPVVEIRQANKVAKEA 165
DB 640 SLDHLETKREKNLQOEISDLTEQIAESAKHIELEKVKQIDQKESLQALAEASGS 699
QY 166 ANRWTDNIPAI-----KSWAKRKEGFEENKIDR 193
DB 700 LEHEGKILRIQLNQNQVSEIDRKIAEKDEIDQ 734

RESULT 26
SMC_METJA STANDARD; PRT; 1169 AA.
AC Q59037;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromosome partition protein smc homolog.
GN MJ1643.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073 (1996).
RN [2]
RP REVISIONS.
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN CHROMOSOME STRUCTURE AND
CC PARTITIONING. ESSENTIAL FOR CHROMOSOME PARTITION (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the SMC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67604; AAB99663.1; --
DR TIGR; MJ1643; --
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR003405; SMC_C.

DR InterPro; IPR003395; SMC_N.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02463; SMC_N; 1.
DR TIGRFAMs; TIGR00650; MG442; 1.
KW Hypothetical protein; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 31 38 ATP (POTENTIAL).
FT DOMAIN 160 521 COILED COIL (POTENTIAL).
FT DOMAIN 673 1032 COILED COIL (POTENTIAL).
SQ SEQUENCE 1169 AA; 136634 MW; B63CCE34E4C03F36 CRC64;

Query Match 10.3%; Score 107.5; DB 1; Length 1169;
Best Local Similarity 25.2%; Pred. No. 10;
Matches 62; Conservative 37; Mismatches 64; Indels 83; Gaps 14;

QY 1 MSKKKGLSAEKKTRMMEIPSETKQVQLKDLKLEKIAK-----EKGITAMSVKVQLSLV 55
DB 787 INELKEYESDENLKRMEIEGE-----LKLKEKAKLNEIDKGLT--VKEIL----- 834
QY 56 DDGMVDCERIGTSNYWAFPSKALHAKHKLVLVLSQSGSKHASLQK-----SIE 108
DB 835 -----IP-----KIEELNKKVSELINKKVKILEKNISFYKESIE 867
QY 109 KAKIGRCETERTP---LAKELSLRDQREQLKAEVE-----KYKQCDQPV--- 151
DB 868 K-NLSILE-EKRRKYBELAKNLKELTEKQELKEIETLERERREILRKVRDIENREL 925
QY 152 -VEEIRQANKVAKAANRW-TDNIFAIKSWAKRK-----FGFEENKIDRTFGIP--- 198
DB 926 MVEKAKYESKLEEEERKLYLCEKVDVSKLEKIDIELEIYIGELENEIKSLSPVNMRAI 985
QY 199 EDFDYI 204
DB 986 EDYNYV 991

RESULT 27
SMC2_HUMAN STANDARD; PRT; 1197 AA.
ID SMC2_HUMAN
AC Q95347; Q9P1P2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Structural maintenance of chromosome 2-like 1 protein (Chromosome-
DE associated protein E) (hCAP-E) (XCAP-E homolog) (PRO0324).
GN SMC2L1 OR SMC2 OR CAPE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH SMC4L1.
RC TISSUE=Testis; Carcinoma;
RX MEDLINE=99007239; PubMed=9789013;
RA Schmiesing J.A., Ball A.R. Jr., Gregson H.C., Alderton J.M., Zhou S.,
RA Yokomori K.;
RT "Identification of two distinct human SMC protein complexes involved
RT in mitotic chromosome dynamics.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12906-12911 (1998).
RN [2]
RP SEQUENCE OF 907-1197 FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Ouyang S., Luo L., Wei H., Zhou G.,
RA Zhou W., Bi J., Zhang Y., Liu M., He F.;
RT "Functional prediction of the coding sequences of 32 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP IDENTIFICATION IN A CONDENSIN COMPLEX WITH SMC4L1 AND CNAP1, AND
RP SUBCELLULAR LOCATION.
RX MEDLINE=20414707; PubMed=10958694;
RA Schmiesing J.A., Gregson H.C., Zhou S., Yokomori K.;
RT "A human condensin complex containing hCAP-C-hCAP-E and CNAP1, a
RT homolog of Xenopus XCAP-D2, colocalizes with phosphorylated histone

H3 during the early stage of mitotic chromosome condensation.";
 RL Mol. Cell. Biol. 20:6996-7006(2000).
 [4]
 RP IDENTIFICATION IN A CONDENSIN COMPLEX WITH SMC4L1; BRRN1; CNAIP1 AND
 CAPG, AND FUNCTION OF THE COMPLEX.
 RX MEDLINE=21101910; PubMed=11136719;
 RA Kimura K., Cuvier O., Hirano T.;
 RT "Chromosome condensation by a human condensin complex in Xenopus egg
 extracts.";
 RL J. Biol. Chem. 276:5417-5420(2001).
 CC -!- FUNCTION: Central component of the condensin complex, a complex
 required for conversion of interphase chromatin into mitotic-like
 CC condense chromosomes. The condensin complex probably introduces
 CC positive supercoils into relaxed DNA in the presence of type I
 CC topoisomerases and converts nicked DNA into positive knotted forms
 CC in the presence of type II topoisomerases.
 CC -!- SUBUNIT: Forms an heterodimer with SMC4L1. Component of the
 CC condensin complex, which contains the SMC2L1 and SMC4L1
 CC heterodimer, and three non SMC subunits that probably regulate the
 CC complex: BRRN1/CAPH, CNAIP1/CAP2 and CAPG.
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. In interphase
 CC cells, the majority of the condensin complex is found in the
 CC cytoplasm, while a minority of the complex is associated with
 CC chromatin. A subpopulation of the complex however remains
 CC associated with chromosome foci in interphase cells. During
 CC mitosis, most of the condensin complex is associated with the
 CC chromatin. At the onset of prophase, the regulatory subunits of
 CC the complex are phosphorylated by CDK2, leading to condensin's
 CC association with chromosome arms and to chromosome condensation.
 CC Dissociation from chromosomes is observed in late telophase.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O95347-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O95347-2; Sequence=VSP_007243, VSP_007244;
 CC Note=No experimental confirmation available;
 CC -!- DOMAIN: The hinge domain, which separates the large intramolecular
 CC coiled coil regions, allows the heterodimerization with SMC4L1,
 CC forming a V-shaped heterodimer (By similarity).
 CC -!- SIMILARITY: Belongs to the SMC family. SMC2 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF092563; AAC72360.1; -;
 DR EMBL; AF113673; AAF29579.1; ALT_INIT.
 DR Genew; HGNC:114011; SMC2L1.
 DR MIM; 605576; -;
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0007001; P:chromosome organization and biogenesis (gen. . .; TAS.
 DR GO; GO:0000067; P:DNA replication and chromosome cycle; TAS.
 DR GO; GO:0007076; P:mitotic chromosome condensation; TAS.
 DR GO; GO:0000070; P:mitotic chromosome segregation; TAS.
 DR InterPro; IPR003405; SMC-C.
 DR InterPro; IPR003395; SMC-N.
 DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF02463; SMC_N; 1.
 KW DNA condensation; Mitosis; Cell cycle; ATP-binding; Coiled coil;
 FT NP BIND 32 39 ATP (POTENTIAL).
 FT DOMAIN 173 507 COILED COIL (POTENTIAL).
 FT DOMAIN 508 671 FLEXIBLE HINGE.
 FT DOMAIN 672 926 COILED COIL (POTENTIAL).
 FT DOMAIN 963 1031 COILED COIL (POTENTIAL).
 FT DOMAIN 1085 1120 ALA/ASP-RICH (DA-BOX).
 FT VARSPIC 1091 1099 SLVALSIL -> OKQNHITG (in isoform 2).
 FT /FTId=VSP_007243.

FT VARSPIC 1100 1197 Missing (in isoform 2).
 FT CONFLICT 907 907 S -> Q (IN REF. 2).
 FT CONFLICT 916 916 H -> N (IN REF. 2).
 FT CONFLICT 998 998 C -> Y (IN REF. 2).
 SQ SEQUENCE 1197 AA; 135780 MW; CC6A7EFA8B34D1B CRC64;
 Query Match 10.3%; Score 107.5; DB 1; Length 1197;
 Best Local Similarity 21.5%; Pred. No. 11;
 Matches 53; Conservative 38; Mismatches 81; Indels 75; Gaps 8;
 QY 7 LSAREKTRMWEIFSETYKDV-----FQKDLKIAKPKGKITAMSV 47
 DB 240 LLAEDTKVRSAELEKEMQDKVFKQBELSNDKKIKALNHEIELEKKDKETGV----- 294
 QY 48 KEVLQSLVDDGMVDCERIGT-SNYWAFPPSKALHARKHLEVSQLESGSQKHSIQKS 106
 DB 295 --ILRSL-EDALAAQRVNTKSQSFADLKKKULACEBSKKELEKNVDSKTLAAKEKE 351
 QY 107 IEKAKIGRCETEETRLAKE-----LSSLDQREQL-----K 138
 DB 352 VKKITDGLHALQEAASNDKDAALAAQHFNAVSAGLSNEDGAETLAGQMACKNDISK 411
 QY 139 AEVE-----YKDCDPQVVEIRQANKVAKAANRWTDNIPAIKSWAKRK 183
 DB 412 AQTEAKQAQMKLKAQQLKQAEVKKMSGYKQDEALEAVKLEKLEA-----EMKK 467
 QY 184 GFPEENK 190
 DB 468 LNYEENK 474
 RESULT 28
 TPML HUMAN
 ID TPML_HUMAN STANDARD; PRT; 284 AA.
 AC P09493; P09494; P10469; Q96IK2; Q9UCY9;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
 GN TPML OR TWSA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE (ISOFORM 1).
 RX MEDLINE=87156701; PubMed=3548719;
 RA Mische S.M., Manjula B.N., Fischetti V.A.;
 RT "Relation of streptococcal M protein with human and rabbit
 RT tropomyosin: the complete amino acid sequence of human cardiac alpha
 RT tropomyosin, a highly conserved contractile protein.";
 RL Biochem. Biophys. Res. Commun. 142:813-818(1987).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Liver;
 RX MEDLINE=88333013; PubMed=3138425;
 RA Colote S., Widada J.S., Ferraz C., Bonhomme F., Marti J.,
 RA Liautard J.-P.;
 RT "Evolution of tropomyosin functional domains: differential splicing
 RT and genomic constraints.";
 RL J. Mol. Evol. 27:228-235(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX TISSUE=Fibroblast;
 RX MEDLINE=88094382; PubMed=3336357;
 RA Lin C.-S., Leavitt J.;
 RT "Cloning and characterization of a cDNA encoding transformation-
 RT sensitive tropomyosin isoform 3 from tumorigenic human fibroblasts.";
 RL Mol. Cell. Biol. 8:160-168(1988).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=88094416; PubMed=3336363;

Query Match 10.2%; Score 107; DB 1; Length 284;
 Best Local Similarity 22.6%; Pred. No. 2.5;
 Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

CC 1 MSKKGLSABEKTRMWEISFETKQVQLKLEKIAPEKGTAMSVKVLQSLVDD--- 57
 CC 73 LAEKATDAE-----ADVASLNRR-OLVVEELDRAQERLATALQKLEAEKADESER 125
 CC 58 GMDVDCRIGTSNYWAPPSPKALHAR-----KHKLVELESQSEGSQKH----- 100
 CC 126 GM-----KVIESRAQKDEERWEIQEIQLEA--KHIAEDADRKYEEVA 166
 CC 101 -----ASLOKSIEKAKI--GRCTEETRLAKELSSLRDOROLKAEVEKYKDCDPQV 151
 CC 167 RKLVIIESDLERAEERAESEGC-----AELEELKVTNNKLSLEAQAEKYSQKEDRY 221
 CC 152 VEEIRQANKVAKAANRWTDNIPAKSWAKRKFGFEENKID 192
 CC 222 EEEIKVLSDKLEAEETR-----AEFAERSVTKLEKSID 254

DR EMEL; X64831; CAA46043.1; --
 DR EMEL; M22479; AAA40483.1; --
 DR PIR; A31380; A60597.
 DR DR SWISS-2DPAGE; P58771; MOUSE.
 DR MGD; MGI:98809; Tpm1.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00261; Tropomyosin; 1.
 DR PRINTS; PS00194; TROPOMYOSIN.
 DR PROSITE; PS00326; TROPOMYOSIN; 1.
 DR Muscle protein; Cytoskeleton; Actin-binding; Coiled coil; Acetylation;
 KW Phosphorylation; Alternative splicing; Multigene family.
 FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 FT MOD RES 283 283 PHOSPHORYLATION (BY SIMILARITY).
 FT VAR5PLIC 258 284 DELVAOKLKYKATSEELDHANDMTSI -> EKVAHAKEEN
 FT LSMHQLDQTLLENNM (in isoform 2).
 FT /FTId=VSP_006580.

CC SEQUENCE 284 AA; 32680 MW; E25609F597A72F4D CRC64;
 CC Query Match 10.2%; Score 107; DB 1; Length 284;
 CC Best Local Similarity 22.6%; Pred. No. 2.5;
 CC Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKGLSABEKTRMWEISFETKQVQLKLEKIAPEKGTAMSVKVLQSLVDD--- 57
 DB 73 LAEKATDAE-----ADVASLNRR-OLVVEELDRAQERLATALQKLEAEKADESER 125
 QY 58 GMDVDCRIGTSNYWAPPSPKALHAR-----KHKLVELESQSEGSQKH----- 100
 DB 126 GM-----KVIESRAQKDEERWEIQEIQLEA--KHIAEDADRKYEEVA 166
 QY 101 -----ASLOKSIEKAKI--GRCTEETRLAKELSSLRDOROLKAEVEKYKDCDPQV 151
 DB 167 RKLVIIESDLERAEERAESEGC-----AELEELKVTNNKLSLEAQAEKYSQKEDRY 221
 QY 152 VEEIRQANKVAKAANRWTDNIPAKSWAKRKFGFEENKID 192
 DB 222 EEEIKVLSDKLEAEETR-----AEFAERSVTKLEKSID 254

RESULT 29
 Tpm1_MOUSE
 ID Tpm1_MOUSE STANDARD; PRT; 284 AA.
 AC P58771; P02558; P19354; P46902; P99034;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
 GN Tpm1 OR Tpm-1 OR TPMA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=A2G; TISSUE=Fast muscle;
 RX MEDLINE=95003687; PubMed=7522680;
 RA Schleef M., Zuehlke C., Schieff F., Jockusch H.;
 RT "Subtractive cDNA cloning as a tool to analyse secondary effects of a
 RT muscle disease. Characterization of affected genes in the myotonic ADR
 RT mouse.";
 RL Neuromuscul. Disord. 4:205-217(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=89219020; PubMed=3244365;
 RA Takenaga K., Nakamura Y., Tokunaga K., Kageyama H., Sakiyama S.;
 RT "Isolation and characterization of a cDNA that encodes mouse
 RT fibroblast tropomyosin isoform 2.";
 RL Mol. Cell. Biol. 8:5561-5565(1988).
 CC -!- FUNCTION: Binds to actin filaments in muscle and nonmuscle cells.
 CC Plays a central role, in association with the troponin complex, in
 CC the calcium dependent regulation of vertebrate striated muscle
 CC contraction. Smooth muscle contraction is regulated by interaction
 CC with caldesmon. In nonmuscle cells is implicated in stabilizing
 CC cytoskeleton actin filaments.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1; Synonyms=Skeletal muscle;
 CC IsoId=P58771-1; Sequence=Displayed;
 CC Name=2; Synonyms=Fibroblast;
 CC IsoId=P58771-2; Sequence=VSP_006580;
 CC -!- DOMAIN: The molecule is in a coiled coil structure. The sequence
 CC exhibits a prominent seven-residues periodicity.
 CC -!- MISCELLANEOUS: THE SEQUENCES OF CARDIAC AND SKELETAL MUSCLES ARE
 CC IDENTICAL.
 CC -!- SIMILARITY: Belongs to the tropomyosin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

CC EMEL; X64831; CAA46043.1; --
 CC EMEL; M22479; AAA40483.1; --
 CC PIR; A31380; A60597.
 CC DR SWISS-2DPAGE; P58771; MOUSE.
 CC MGD; MGI:98809; Tpm1.
 CC InterPro; IPR000533; Tropomyosin.
 CC Pfam; PF00261; Tropomyosin; 1.
 CC PRINTS; PS00194; TROPOMYOSIN.
 CC PROSITE; PS00326; TROPOMYOSIN; 1.
 CC Muscle protein; Cytoskeleton; Actin-binding; Coiled coil; Acetylation;
 KW Phosphorylation; Alternative splicing; Multigene family.
 FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).
 FT MOD RES 283 283 PHOSPHORYLATION (BY SIMILARITY).
 FT VAR5PLIC 258 284 DELVAOKLKYKATSEELDHANDMTSI -> EKVAHAKEEN
 FT LSMHQLDQTLLENNM (in isoform 2).
 FT /FTId=VSP_006580.

CC SEQUENCE 284 AA; 32680 MW; E25609F597A72F4D CRC64;
 CC Query Match 10.2%; Score 107; DB 1; Length 284;
 CC Best Local Similarity 22.6%; Pred. No. 2.5;
 CC Matches 50; Conservative 39; Mismatches 64; Indels 68; Gaps 10;

QY 1 MSKKGLSABEKTRMWEISFETKQVQLKLEKIAPEKGTAMSVKVLQSLVDD--- 57
 DB 73 LAEKATDAE-----ADVASLNRR-OLVVEELDRAQERLATALQKLEAEKADESER 125
 QY 58 GMDVDCRIGTSNYWAPPSPKALHAR-----KHKLVELESQSEGSQKH----- 100
 DB 126 GM-----KVIESRAQKDEERWEIQEIQLEA--KHIAEDADRKYEEVA 166
 QY 101 -----ASLOKSIEKAKI--GRCTEETRLAKELSSLRDOROLKAEVEKYKDCDPQV 151
 DB 167 RKLVIIESDLERAEERAESEGC-----AELEELKVTNNKLSLEAQAEKYSQKEDRY 221
 QY 152 VEEIRQANKVAKAANRWTDNIPAKSWAKRKFGFEENKID 192
 DB 222 EEEIKVLSDKLEAEETR-----AEFAERSVTKLEKSID 254

RESULT 30
 Tpm1_RABIT
 ID Tpm1_RABIT STANDARD; PRT; 284 AA.
 AC P58772; P02558; P46902; P99034;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tropomyosin 1 alpha chain (Alpha-tropomyosin).
 GN Tpm1 OR TPMA.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Zealand white; TISSUE=Skeletal muscle;
 RX MEDLINE=95348266; PubMed=7622625;
 RA Kluebe L., Maeda K., Miesgel A., Fujita-Becker S., Maeda Y.,
 RA Taibo G., Houthaave T., Kellner R.;
 RT "Rabbit skeletal muscle alpha alpha-tropomyosin expressed in
 RT baculovirus-infected insect cells possesses the authentic N-terminus
 RT structure and functions.";
 RL J. Muscle Res. Cell Motil. 16:103-110(1995).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=78109457; PubMed=624724;
 RA Stone D., Smillie L.B.;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2004, 08:36:32 ; Search time 40 Seconds

(without alignments)
1617.031 Million cell updates/sec

Title: US-10-087-190-3

Perfect score: 1047

Sequence: 1 MSKKGLSAEKRTRMEIF.....FEENKIDRTFGPEDFDVID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPREMBL.25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1047	100.0	205	4	Q9BWT6 homo sapien
2	954	91.1	205	11	Q8K396 mus musculus
3	948	90.5	205	11	Q9D0A1 mus musculus
4	462	44.1	196	5	Q86E28 schistosoma
5	440	42.0	230	10	Q8GYD2 arabisidopsis
6	340.5	32.5	203	5	Q8SUA9 encephalito
7	197	18.8	128	10	Q8SZE5 arabisidopsis
8	136.5	13.0	910	13	Q7T2F8 brachydanio
9	126	12.0	609	17	Q8TXA4 methanopyru
10	122	11.7	2093	5	Q9N435 caenorhabdi
11	122	11.7	10578	5	Q818F5 caenorhabdi
12	122	11.7	18519	5	Q818F6 caenorhabdi
13	122	11.7	18534	5	Q818F7 caenorhabdi
14	120.5	11.5	172	16	Q824V3 chlamydophi
15	120.5	11.5	448	2	Q84H59 anaplasma m
16	120.5	11.5	1732	5	Q9VJ35 drosophila

17	120	11.5	976	11	Q99MI2
18	120	11.5	1120	11	Q99MI1
19	118.5	11.3	1177	1	Q87711 pyrococcus
20	118.5	11.3	1291	17	Q8TZV2
21	117	11.2	768	10	Q80951 arabisidopsis
22	117	11.2	795	4	Q9H237
23	117	11.2	948	11	Q8CIY9
24	117	11.2	948	11	Q811U3
25	116.5	11.1	920	2	Q84HV4
26	116	11.1	559	5	Q24788
27	115.5	11.0	269	17	Q8TQ04
28	115	11.0	284	13	Q805D2
29	114.5	10.9	365	4	Q96N90
30	114.5	10.9	448	2	Q84H67
31	114.5	10.9	478	11	Q8CG58
32	114.5	10.9	788	4	Q86TF6
33	114	10.9	764	10	Q49371
34	114	10.9	783	10	Q8H1E5
35	113.5	10.8	415	11	Q8CG53
36	113.5	10.8	450	11	Q8CG57
37	113.5	10.8	450	11	Q8CG56
38	113.5	10.8	473	11	Q8CG52
39	113	10.8	284	13	Q805C8
40	113	10.8	720	4	Q8IUD4
41	113	10.8	783	10	Q9C5L5
42	113	10.8	948	4	Q8UIK7
43	113	10.8	992	4	Q8IUD5
44	113	10.8	1003	4	Q9UP81
45	113	10.8	1938	6	Q8MJV0
46	112.5	10.7	1391	11	Q922J3
47	112	10.7	284	13	Q87349
48	112	10.7	284	13	Q8AV86
49	112	10.7	559	5	Q24796
50	112	10.7	1945	6	Q97757
51	111.5	10.6	284	13	Q805C7
52	111.5	10.6	502	11	Q8CG54
53	111.5	10.6	1327	11	Q61595
54	111	10.6	284	5	Q02389
55	111	10.6	284	13	Q8JIM7
56	111	10.6	284	13	Q90WE7
57	111	10.6	354	5	Q92122
58	111	10.6	559	5	Q05768
59	110.5	10.6	1320	11	Q9JK25
60	110.5	10.6	293	13	Q805D0
61	110.5	10.6	308	4	Q9Y427
62	110.5	10.6	451	11	Q8BHM8
63	110.5	10.6	629	1	Q07116
64	110.5	10.6	629	17	Q9HEM7
65	110.5	10.6	671	13	Q9YHD4
66	110	10.5	284	13	Q91490
67	110	10.5	284	13	Q7SZ31
68	110	10.5	1088	4	Q8IUD3
69	110	10.5	1116	4	Q8IUD2
70	109.5	10.5	168	16	Q9Z7J9
71	109.5	10.5	284	13	Q87348
72	109.5	10.5	400	16	Q66577
73	109.5	10.5	1298	5	Q44199
74	109.5	10.5	1663	4	Q8WZ74
75	109	10.4	284	13	Q90236
76	109	10.4	284	13	Q8JIM8
77	109	10.4	284	13	Q805C6
78	109	10.4	284	13	Q805C3
79	109	10.4	284	13	Q7ZVK9
80	109	10.4	699	13	Q98T11
81	109	10.4	1937	6	Q9TV62
82	108.5	10.4	600	11	Q9EPM6
83	108.5	10.4	712	11	Q8BIJ7
84	108	10.3	451	16	Q8XJ77
85	108	10.3	670	10	Q82351
86	108	10.3	725	10	Q8VYU6
87	108	10.3	811	4	Q8NE23
88	107.5	10.3	248	13	Q7T3F0
89	107.5	10.3	1085	10	Q9CA42

90 107.5 10.3 1137 5 Q93250 Q93250 caenorhabdi
 91 107 10.2 123 17 Q973R1 Q973R1 sulfolobus
 92 107 10.2 280 11 P70524 P70524 rattus norv
 93 107 10.2 284 11 Q8BSH3 Q8BSH3 mus musculu
 94 107 10.2 284 13 Q9DEB7 Q9DEB7 pennaria ar
 95 107 10.2 284 13 Q805C5 Q805C5 fugu rubrip
 96 107 10.2 287 11 Q91XN6 Q91XN6 rattus norv
 97 107 10.2 879 13 Q91HD8 Q91HD8 rana catesb
 98 107 10.2 1177 16 Q8RCY8 Q8RCY8 thermoanaer
 99 107 10.2 1743 5 Q96063 Q96063 dugesia jap
 100 107 10.2 1941 4 Q86T56 Q86T56 homo sapien

ALIGNMENTS

RESULT 1
 Q9BWT6 PRELIMINARY; PRT; 205 AA.
 AC Q9BWT6
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GAJ.
 GN GAJ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Solis G., Hofer H.W.;
 RT "GAJ protein isolated from Jurkat cells."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY028916; AAK26168.1; -.
 DR EMBL; BC032142; AAH32142.1; -.
 DR IncerPro; IPR005647; Mnd1.
 DR Pfam; PF03962; Mnd1; 1.
 SQ SEQUENCE 205 AA; 23753 MW; 95B0B14068DA0B51 CRC64;

Query Match 100.0%; Score 1047; DB 4; Length 205;
 Best Local Similarity 100.0%; Pred. No. 1.9e-64;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKKKGLSABEKTRMVEIFSETKDVFLKLEKIAPEKIGITAMSVKEVLSLVDGMY 60
 DB 1 MSKKKGLSABEKTRMVEIFSETKDVFLKLEKIAPEKIGITAMSVKEVLSLVDGMY 60
 QY 61 DCRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKASLOKSIKAKIGRCETEER 120
 DB 61 DCRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKASLOKSIKAKIGRCETEER 120
 QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAIKSWA 180
 DB 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAIKSWA 180
 QY 181 KRKFGPEENKIDRTFGIPEDFYID 205
 DB 181 KRKFGPEENKIDRTFGIPEDFYID 205

RESULT 2
 Q8K396 PRELIMINARY; PRT; 205 AA.
 AC Q8K396
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE RIKEN CDNA 2610034E18 gene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC027741; AAH27741.1; -.
 DR InterPro; IPR005647; Mnd1.
 DR Pfam; PF03962; Mnd1; 1.
 SQ SEQUENCE 205 AA; 23849 MW; 122C3FA9E4325120 CRC64;

Query Match 91.1%; Score 954; DB 11; Length 205;
 Best Local Similarity 89.8%; Pred. No. 4.7e-56;
 Matches 184; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
 QY 1 MSKKKGLSABEKTRMVEIFSETKDVFLKLEKIAPEKIGITAMSVKEVLSLVDGMY 60
 DB 1 MSKKKGLSABEKTRMVEIFSETKDVFLKLEKIAPEKIGITAMSVKEVLSLVDGMY 60
 QY 61 DCRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKASLOKSIKAKIGRCETEER 120
 DB 61 DCRIGTSNYWAPPSPKALHARKHKLVELESQSEGSQKASLOKSIKAKIGRCETEER 120
 QY 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAIKSWA 180
 DB 121 TRLAKELSSLRDQREQLKAEVEKYKDCDPQVVEIRQANKVAKAANRWTDNIFAIKSWA 180
 QY 181 KRKFGPEENKIDRTFGIPEDFYID 205
 DB 181 KRKFGPEENKIDRTFGIPEDFYID 205

RESULT 3
 Q9D0A1 PRELIMINARY; PRT; 205 AA.
 AC Q9D0A1
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE 2610034E18Rik protein.
 GN 2610034E18RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 403:685-690(2001).
 EMBL; AK011664; BAB27765.1; -.


```
DR MGD; MGI:1924165; 2610034E18Rik.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 205 AA; 23909 MW; 09368E19E4224021 CRC64;

Query Match          90.5%; Score 948; DB 11; Length 205;
Best Local Similarity 89.3%; Pred. No. 1.2e-57;
Matches 183; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSKKKGLSAEKRTRMWEIFSETKDVQFKDLEKAPKEKGITAMSVKEVLQSLVDDGMV 60
DB 1 MSKKRGLSGEKRTRMWEIFFETKDVQFKDLEKAPKEKGITAMSVKEVLQSLVDDGMV 60

QY 61 DCEIRGTSTNYWAFPSKALHARKHKLVEVLSOLSGSOKHASLOKSIKAKIGRCETEER 120
DB 61 DCEIRGTSTNYWAFPSKALHARKHKLVEVLSOLSGSOKHASLOKSIKAKIGRCETEER 120

QY 121 TLAKELSLRQREQLKAEVKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
DB 121 AMLAKELFSFRQRLKAEVKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180

QY 181 KRKFGPEENKIDRTGIPEDFDYID 205
DB 181 KRKFGPEESKIDNFGIPEDFDYID 205

RESULT 4
Q86EZ8 PRELIMINARY; PRT; 196 AA.
AC Q86EZ8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Clone ZD1259 mRNA sequence
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
[1]
RP SEQUENCE FROM N.A.
RA Liu F., Hu W., Yan Q., Xu X., Zhu Z., Zhang X., Wang J., Sun J.,
RA Xu X., Wang Z., Zeng L., Rong Y., Wu X., Qu J., Xu J., Huang J.,
RA Ma Y., Wang S., Wang Z., Xue C., Feng Z., Chen Z., Han Z.;
RT "The full-length cDNA of S. japonicum genes."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY223066; AAP06089.1; -.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
SQ SEQUENCE 196 AA; 23163 MW; B30F6F088D7123F0 CRC64;

Query Match          44.1%; Score 462; DB 5; Length 196;
Best Local Similarity 47.4%; Pred. No. 2.4e-24;
Matches 92; Conservative 34; Mismatches 68; Indels 0; Gaps 0;

QY 11 EKRTMWEIFSETKDVQFKDLEKAPKEKGITAMSVKEVLQSLVDDGMVDCERIGTSTNY 70
DB 2 KSRQRMWDFYKDFKQFFQKLERLCQKEKNGINSKVDVLSVHVDGLVDTKIGTSTNY 61

QY 71 YWAFPSKALHARKHKLVEVLSOLSGSOKHASLOKSIKAKIGRCETEERTLAKELSSL 130
DB 62 FWAFFSKAQAQLNNEIKYTGIDHTRNQIFKTRSLNEALSKRDXTEERNRIINELTEL 121

QY 131 RQREQLKAEVKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWAKRKGFGFENK 190
DB 122 KILLESITAELOPLEKHPDRLSLRFQQLVALDSANRWTDNIFIKSWLSNKFSLDEAT 181

QY 191 IDRTGIPEDFDYID 204
DB 182 FCRQFIPENFDYI 195

RESULT 5
Q8GYD2 PRELIMINARY; PRT; 230 AA.
AC Q8GYD2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN AtG29170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA."
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK117713; BAC42364.1; -.
DR EMBL; BT005435; AAC63855.1; -.
DR InterPro; IPR005647; Mnd1.
DR Pfam; PF03962; Mnd1; 1.
KW Hypothetical protein.
SQ SEQUENCE 230 AA; 26402 MW; 131B5146D8C91ED CRC64;

Query Match          42.0%; Score 440; DB 10; Length 230;
Best Local Similarity 43.8%; Pred. No. 9.4e-23;
Matches 91; Conservative 43; Mismatches 70; Indels 4; Gaps 2;

QY 1 MSKKKGLSAEKRTRMWEIFSETKDVQFKDLEKAPKEKGITAMSVKEVLQSLVDDGMV 60
DB 1 MSKKRGLSGEKRTRMWEIFFETKDVQFKDLEKAPKEKGITAMSVKEVLQSLVDDDLV 59

QY 61 DCEIRGTSTNYWAFPSKALHARKHKLVEVLSOLSGSOKHASLOKSIKAKIGRCETEER 120
DB 60 AKDXIGISYFWSLPSCAGNQLRSVRQKLESDDLQGSNKLAEVLDOCEALKKGRSEER 119

QY 121 TLAKELSLRQREQLKAEVKYKDCDPQVVEIRQANKVAKAEANRWTDNIFAIKSWA 180
DB 120 TEALTQLKDIEKKHKDKNEMVQFADNDPATLEAKRNAIEVAHQSANRWTDNIFTLRQC 179

QY 181 KRKFGPEENKIDRTF---GIPEDFDYID 205
DB 180 SNNFPQAKQLEHLYTEAGITDFDYIE 207

RESULT 6
Q8SUA9 PRELIMINARY; PRT; 203 AA.
AC Q8SUA9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein ECU10_1600.
GN ECU10_1600.
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
[1]
RP SEQUENCE FROM N.A.
```



```
FT NON_TER 910 910
SQ SEQUENCE 910 AA; 101651 MW; DA4413EC5C743070 CRC64;

Query Match
Best Local Similarity 13.0%; Score 136.5; DB 13; Length 910;
Matches 48; Conservative 47; Mismatches 104; Indels 15; Gaps 5;

QY 2 SKKGLSAEKKTRMWEIFSETKDVFLQKDLKIAPEKKGITAMSVKVLQSLVDDGMVD 61
DB 228 AKKEAREALEAKDRYMEVADTADIAEMATLQKMAEERAEASLQDETSLKEKLEELTMD 287
QY 62 CER1-----GTSNYWAPPKALHAR-KHKLEVLSEQLSEGSOKHASLOKSIEKA 110
DB 288 MILKHEIEKSGDGAASSYHVQLEEQNARLKEALVPMRDLSSSEKQEHVKLQKQMEK- 346
QY 111 KIGRCET--BERTRIAKELSSLRDQREQLKAEVYKDCDPQVVEIRQANKVAKAANR 168
DB 347 KNSLETLSRQKLEQENKLAEDTIDELKEQVDAALGAE-EMVETLTFRNLDLBEKRV 405
QY 169 WTDNIFAIKSWAKRKGFEENKIDRTFGIPEDFD 202
DB 406 LRETVSDLESINEMDELQENARETELEUREQLD 439

RESULT 9
Q8TXA4
ID Q8TXA4 PRELIMINARY; PRT; 609 AA.
AC Q8TXA4
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE UNcharacterized protein.
GN MK0771.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaia K.V., Makarova G.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AB010369; RA001985.1; -.
DR HSSP; P04268; 1IC2.
DR InterPro; IPR002017; Spectrin.
KW Complete proteome.
SQ SEQUENCE 609 AA; 69552 MW; AB10C9780DC5AD78 CRC64;

Query Match
Best Local Similarity 12.0%; Score 126; DB 17; Length 609;
Matches 46; Conservative 30; Mismatches 45; Indels 26; Gaps 7;

QY 3 KKKGLSAEKKTRMWEIFSETKDVFLQKDLKIAPEKKGITAMSVKVLQSLVDDGMVD 61
DB 217 KEKYNIEKERDLKE---ETKEVGKLDQLAKLSQK-----LKEV-KSERDDLANE 264
QY 62 CERIGTSNYWAPPKALHARKHKLEVLSEQLSEGSOKHASLOKSIEKAK--IGRCETEE 119
DB 265 VEALRNE-----EKLKKIKDKLSLSNLQKLDKREKKLEKARQHQIGKLR-BE 313
QY 120 RTRIAKELSSLRDQREQLKAEVYKDCD 146
DB 314 IKRDEIRKLRKAQSKLDEIKRYEE 340

RESULT 10
Q9N435
ID Q8ISF5 PRELIMINARY; PRT; 10578 AA.
```

```
Q9N435 PRELIMINARY; PRT; 2083 AA.
Q9N435;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y385A.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Sun H., Geisel C.;
RT "The sequence of C. elegans cosmid Y385A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024760; AAF59460.2; -.
DR WormPep; Y385A.1; CE26070.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00041; fn3; 2.
DR PRINTS; PR00014; ENTPEIII.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00409; IG; 1.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 2083 AA; 231174 MW; 9270303ADE2D7C2F CRC64;

Query Match
Best Local Similarity 11.7%; Score 122; DB 5; Length 2083;
Matches 59; Conservative 44; Mismatches 87; Indels 58; Gaps 9;

QY 3 KKKGLSAEKKTRMWEIFSETKDVFLQKDLKIAPEK--EKGITAMSVKVL-QSLVDDG- 58
DB 1219 KQTGLEKDDKSTK----DSSSKETVDEKPKKVKLKTKESDSSISQKSVTSKTVVSSGG 1274
QY 59 -----MVDCE-----IGTSNYWAPPKALHARK 83
DB 1275 PSESTQKVADAARKQKQETDEKQKLEAEITAKKSADSKLEAKSLKAAAEVZAQKKQ 1334
QY 84 HKLEVLSEQLSEGSOKHASLOKSIEK-AKIGRC-----ETEERTRIAKELSSLRD 132
DB 1335 EKDEQLKLDTEAASKKAAAEKLEKQKQAIKKAADAVAKKEKELABKQKLESAATKKA 1394
QY 133 QREOLKAEVYKDCDPQVVEIRQANKVAKAANRWTNIFAIKSWAKRKGFEENKID 192
DB 1395 AAEKLEQEKQKDAETASTIEKQEKQKLEAQEQSKLEVD---AKKSAEKQKLE-SETKSK 1450
QY 193 RTFGIPED 200
DB 1451 KTEAPKE 1458

RESULT 11
Q8ISF5 PRELIMINARY; PRT; 10578 AA.
ID Q8ISF5
```

```
AC Q8ISF5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1kDa.1 protein.
GN ISOF.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP MEDLINE=22269627; PubMed=12381307;
RX Flaherty D., Gernert K., Shmeleva N., Tang X., Mercer K.,
RA Borodovsky M., Benian G.;
RT "Titins in Caenorhabditis elegans with Unusual Features: Coiled-coil
RT Domains, Novel Regulation of Kinase Activity and Two New Possible
RT Elastic Regions.";
RL J. Mol. Biol. 323:533-549(2002).
DR EMBL, AY130758; AAN61519.1; -.
DR PIR, E89066; E89066.
DR PIR, T33247; T33247.
DR GO, GO:000524; F:ATP binding; IEA.
DR GO, GO:0008237; F:metalloproteinase activity; IEA.
DR GO, GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO, GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO, GO:0008270; F:zinc ion binding; IEA.
DR GO, GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO, GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro, IPR008938; ARM.
DR InterPro, IPR001064; Crystallin.
DR InterPro, IPR003961; FN III.
DR InterPro, IPR008957; FN III-like.
DR InterPro, IPR001023; Hsp70.
DR InterPro, IPR007110; Ig-like.
DR InterPro, IPR003598; Ig_c2.
DR InterPro, IPR003596; Ig_v.
DR InterPro, IPR006025; Pept_M_Zn_BS.
DR InterPro, IPR000719; Prot_kinase.
DR InterPro, IPR002290; Ser_thr_pkinase.
DR InterPro, IPR008271; Ser_thr_pkin_AS.
DR InterPro, IPR001245; Tyr_pkinase.
DR Pfam, PF00041; fn3; 11.
DR Pfam, PF00047; ig; 43.
DR Pfam, PF00069; pkinase; 1.
DR ProDom, PD000089; Hsp70; 1.
DR ProDom, PD000001; Prot_kinase; 1.
DR SMART, SMC0060; FN3; 11.
DR SMART, SMC0408; IGc2; 37.
DR SMART, SMC0406; IGv; 4.
DR SMART, SMC0220; S_TKC; 1.
DR SMART, SMC0219; Tyrc; 1.
DR PROSITE, PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE, PS00835; IG LIKE; 38.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE, PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 18519 AA; 2051869 MW; 6A8441C5D0BA7729 CRC64;

Query Match 11.7%; Score 122; DB 5; Length 18519;
Best Local Similarity 23.8%; Pred. No. 78;
Matches 59; Conservative 44; Mismatches 87; Indels 58; Gaps 9;

QY 3 KKGGLSAEERTRMEIFSETKDVFLKLEKIAPK--EKGITAMSVKEVL--QSLVDDG- 58
DB 9739 KQTGLEKDDKSTK-----DSESEKTVDEPKPKVKLKKTEKSDSSISQKSVTKTVVESGG 9794

QY 59 -----MVDCER-----IGTSNYWAFPSKALHARK 83
DB 9795 PSESETQKVADAAKQKQETDEKQKLEAEITAKSADEKSKLEAEKSLKAAAEVAAKKQK 9854

QY 84 HKLEVLSQLSEGSQKSHASLOKSIK-AKIGRC-----ETERTRLAKELSSLRD 132
DB 9855 EKDEQLKLDTEAAASKAAAEKLEKLEKQAQIKKAAAEADAVKKEKLEAEKQKLESEAAATKA 9914

QY 133 QREOLKAEVEKYKDCDPQVVEEIRQANKVAKAANWTNIPAIKSWAKRKFGEENKID 192
DB 9915 AAEKLEBKQKDAETASTIEKQKEKLEAQEQSKLEVD---AKSAEKQKLE-SETKSK 9970

QY 193 RTFGIPED 200
DB 9971 KTEAPKE 9978

RESULT 12
Q8ISF6
ID Q8ISF6 PRELIMINARY; PRT; 18519 AA.
AC Q8ISF6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 2kDa.2 protein.
GN ISOF.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
```

```
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22269627; PubMed=12381307;
RA Flaherty D., Gernert K., Shmeleva N., Tang X., Mercer K.,
RA Borodovsky M., Benian G.;
RT "Titins in Caenorhabditis elegans with Unusual Features: Coiled-coil
RT Domains, Novel Regulation of Kinase Activity and Two New Possible
RT Elastic Regions.";
RL J. Mol. Biol. 323:533-549(2002).
DR EMBL, AY130758; AAN61518.1; -.
DR PIR, E89066; E89066.
DR PIR, T33247; T33247.
DR GO, GO:000524; F:ATP binding; IEA.
DR GO, GO:0008237; F:metalloproteinase activity; IEA.
DR GO, GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO, GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO, GO:0008270; F:zinc ion binding; IEA.
DR GO, GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO, GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro, IPR008938; ARM.
DR InterPro, IPR001064; Crystallin.
DR InterPro, IPR003961; FN III.
DR InterPro, IPR008957; FN III-like.
DR InterPro, IPR001023; Hsp70.
DR InterPro, IPR007110; Ig-like.
DR InterPro, IPR003598; Ig_c2.
DR InterPro, IPR003596; Ig_v.
DR InterPro, IPR006025; Pept_M_Zn_BS.
DR InterPro, IPR000719; Prot_kinase.
DR InterPro, IPR002290; Ser_thr_pkinase.
DR InterPro, IPR008271; Ser_thr_pkin_AS.
DR InterPro, IPR001245; Tyr_pkinase.
DR Pfam, PF00041; fn3; 11.
DR Pfam, PF00047; ig; 43.
DR Pfam, PF00069; pkinase; 1.
DR ProDom, PD000089; Hsp70; 1.
DR ProDom, PD000001; Prot_kinase; 1.
DR SMART, SMC0060; FN3; 11.
DR SMART, SMC0408; IGc2; 37.
DR SMART, SMC0406; IGv; 4.
DR SMART, SMC0220; S_TKC; 1.
DR SMART, SMC0219; Tyrc; 1.
DR PROSITE, PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE, PS00835; IG LIKE; 38.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE, PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 18519 AA; 2051869 MW; 6A8441C5D0BA7729 CRC64;

Query Match 11.7%; Score 122; DB 5; Length 18519;
Best Local Similarity 23.8%; Pred. No. 78;
Matches 59; Conservative 44; Mismatches 87; Indels 58; Gaps 9;

QY 3 KKGGLSAEERTRMEIFSETKDVFLKLEKIAPK--EKGITAMSVKEVL--QSLVDDG- 58
DB 9759 KQTGLEKDDKSTK-----DSESEKTVDEPKPKVKLKKTEKSDSSISQKSVTKTVVESGG 9814

QY 59 -----MVDCER-----IGTSNYWAFPSKALHARK 83
DB 9815 PSESETQKVADAAKQKQETDEKQKLEAEITAKSADEKSKLEAEKSLKAAAEVAAKKQK 9874

QY 84 HKLEVLSQLSEGSQKSHASLOKSIK-AKIGRC-----ETERTRLAKELSSLRD 132
DB 9875 EKDEQLKLDTEAAASKAAAEKLEKLEKQAQIKKAAAEADAVKKEKLEAEKQKLESEAAATKA 9934

QY 133 QREOLKAEVEKYKDCDPQVVEEIRQANKVAKAANWTNIPAIKSWAKRKFGEENKID 192
DB 9935 AAEKLEBKQKDAETASTIEKQKEKLEAQEQSKLEVD---AKSAEKQKLE-SETKSK 9990

QY 193 RTFGIPED 200
DB 9971 KTEAPKE 9978
```

```
Db 9991 KTEEPAPKE 9998

RESULT 13
Q81SF7 PRELIMINARY; PRT; 18534 AA.
AC Q81SF7;
DT 01-MAR-2003 (T-REMBLrel. 23, Created)
DT 01-MAR-2003 (T-REMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE 2XDA_1 protein.
GN ISOF.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae.
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2269627; PubMed=12381307;
RA Flaherty D., Gernert K., Shmeleva N., Tang X., Mercer K.,
RA Borodovsky M., Benian G.;
RT "Titins in Caenorhabditis elegans with Unusual Features: Coiled-coil
RT Domains, Novel Regulation of Kinase Activity and Two New Possible
RT Elastic Regions."
RL J. Mol. Biol. 323:533-549(2002).
DR EMBL; AX130758; AAN61517.1; -.
DR PIR; E89066; E89066.
DR PIR; T33247; T33247.
DR GO; GO:0005524; F-ATP binding; IEA.
DR GO; GO:0008237; F-metalloproteinase activity; IEA.
DR GO; GO:0004574; F-protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F-protein-tyrosine kinase activity; IEA.
DR GO; GO:0008270; F-zinc ion binding; IEA.
DR GO; GO:0006468; P-protein amino acid phosphorylation; IEA.
DR GO; GO:0006508; P-proteinolysis and peptidolysis; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR003951; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR001023; Hsp70.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR008271; Ser_Thr_pkinase.
DR Pfam; PF00041; fn3; 11.
DR Pfam; PF00047; ig; 43.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000089; Hsp70; 1.
DR SMART; SM00001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 11.
DR SMART; SM00408; IGC2; 37.
DR SMART; SM00406; IGV; 4.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TY_KC; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS00835; IG_LIKE; 38.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 18534 AA; 2053504 MW; BC8A682B943C8C0A CRC64;

Query Match 11.7%; Score 122; DB 5; Length 18534;
Best Local Similarity 23.8%; Pred. No. 78;
Matches 59; Conservative 44; Mismatches 87; Indels 58; Gaps 9;

QY 3 KKXGLAEEKRTMWEIFSTKDFQKDLKAPK--EKGITAMSVKEVLQSLVDDG--58
Db 9759 KQTGLEKDKSTK-----DSSEKTVDEKPKKVLKKKTKSDSSISQKSVTSKTVESGG 9814
```

```
QY 59 -----MVDCE-----IGTSNYWAPPSKALHARK 83
Db 9815 PSESTQKVADARKQKETDEKQKLEAEITAKKSADESKLEAESKLKAAEVAEAKKQK 9874
QY 84 HKLEVLSEQLSESGQKHAFLQKSIK-AKIGRC-----ETERTLAKELSLRD 132
Db 9875 EKDEQLKLDTEAASKAAAEKLEKQKQIKKAAEADAVKKEKLEAKQKLESAAYTKA 9934
QY 133 QREOLKAEVEKYKDCDPQVVEITRQANKVAEAAANRWTDNIFAIKSWAKKFGFEENKID 192
Db 9935 AAELKLEEQKKDAETASTIEKQEQKLAQEQSKLEVD---AKKSAEKQKLE-SEYKSK 9990
QY 193 RTFGIPED 200
Db 9991 KTEEPAPKE 9998

RESULT 14
Q824V3 PRELIMINARY; PRT; 172 AA.
AC Q824V3;
DT 01-JUN-2003 (T-REMBLrel. 24, Created)
DT 01-JUN-2003 (T-REMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00036.
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiophila.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzaple E., Khouri H., Federova N.B., Carty H.A.,
RA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavolil P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016994; AAP04788.1; -.
DR TIGR; CCA00036; -.
KW Hypothetical protein; Complete proteome
SQ SEQUENCE 172 AA; 20429 MW; 5F00B0478E6F8311 CRC64;

Query Match 11.5%; Score 120.5; DB 16; Length 172;
Best Local Similarity 27.1%; Pred. No. 0.56;
Matches 46; Conservative 38; Mismatches 53; Indels 33; Gaps 8;

QY 10 EKEKTRMWEIFSETKDYFQKDLKAPKKEGITAMSVKEVLQSLVDDGMDVCERIGTSN 69
Db 25 KEKR-RLLEIEQE-----KLREIARDKVKNHVMQKIQQ-LRELDGTT-----68
QY 70 YWAPPSKALHARKHKLEVLSEQLSE-----GSQKHAFL--QKSIKAKIG--RCETEE 119
Db 69 -----SDAVLQKAYIKVAVQLAEEEEKYKQKESVLAASKLEKAEVNLAKRKEE 122
QY 120 RTRLAKELSLRDQREQLKAEVEKYKDCDPQVVEITRQANKVAEAAANRW 169
Db 123 KTRLHKE-ENMKALKEAREVEKEQDEMGMQLLQLRKNKQKQESGESSW 171

RESULT 15.
Q84H59 PRELIMINARY; PRT; 448 AA.
AC Q84H59;
DT 01-JUN-2003 (T-REMBLrel. 24, Created)
DT 01-JUN-2003 (T-REMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)
DE Major surface protein 3 (Fragment).
```

RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Ceinikher S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blasej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavanyi A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davis P.,
RA de laRiba B., Delecher A., Deng Z., Meyers A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov S.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan K.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

```

QY 101 ASLQK---SIEKAKIGRCETEERTRLAKE--LSSLRDREQLKAEV--EKYKDCD----- 148
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1137 AAVERRQALEXERLILKLEKMNWETRLKEQRIKGKQCEKQKQALAREKARDREERLLA 1196
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 149 -----POVVEEI-RQANKVAKAEANRWTDNIPAIKSWA 180
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1197 LQVQOQQTTEELQKILQKQESARHEENIEHIRQRA 1234
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 17
Q99MI2 PRELIMINARY; PRT; 976 AA.
AC Q99MI2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Rab6-interacting protein 2 isoform A.
GN RAB6IP2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Monier S., Jancuex-Lerosey I., Jollivet F., Goud B.;
RT "Characterization of a novel interaction partner of the small GTPase
Rab6."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF340028; AAK26381.1; -.
DR MGD; MGI:2151013; Rab6IP2.
DR InterPro; IPR002017; Spectrin.
SQ SEQUENCE 976 AA; 111931 MW; 48731867C8D8CD6F CRC64;

Query Match 11.5%; Score 120; DB 11; Length 976;
Best Local Similarity 24.2%; Pred. No. 4.1;
Matches 46; Conservative 32; Mismatches 72; Indels 40; Gaps 6;

QY 11 EKRTMWEIFSETKQVFQKDL-EKIA-----PKEKGITAMSVKEVLQSLVDGMDVDCER 64
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 639 DEREKQEEIDTYKD---LKDREKVSLLQGLDSEKASLLDIKEHASSLASSGLKKDSR 695
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 65 IGTSNYWAFPSKALHARKHKLVEVLSQSEGSKHASLQKSIKAKIGRCETEERTRLA 124
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 696 LKTLLEI-----ALEQKKECLKMSQLK-----KAHEATLEA 727
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 125 KELSLRDREQLKAEVEKYKCDPQVVEIRQANKVAKAEANRWTD---NIPAIKSWAK 181
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 728 RASPEMSDRIQQLEREISRYKDESSKAQTEVDRLLLEILKEVENEKNDKDKKIAELESLS 787
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 182 RKFGFEENKI 191
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 788 RQVKDQNKV 797
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 18
Q99MI1 PRELIMINARY; PRT; 1120 AA.
AC Q99MI1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rab6-interacting protein 2 isoform B (RAB6IP2).
GN RAB6IP2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Monier S., Jancuex-Lerosey I., Jollivet F., Goud B.;
RT "Characterization of a novel interaction partner of the small GTPase
Rab6."

```

```

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Inoue E., Deruchi-Tawarada M., Takao-Rikitsu E., Inoue M., Ohtsuka T.;
RT "CASTbeta (JUN-2003) to the EMBL/GenBank/DBJ databases."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF340029; AAK26382.1; -.
DR EMBL; AY316692; AAP83581.1; -.
DR MGD; MGI:2151013; Rab6IP2.
DR InterPro; IPR002017; Spectrin.
SQ SEQUENCE 1120 AA; 128330 MW; A542B526FAEDF9C7 CRC64;

Query Match 11.5%; Score 120; DB 11; Length 1120;
Best Local Similarity 24.2%; Pred. No. 4.8;
Matches 46; Conservative 32; Mismatches 72; Indels 40; Gaps 6;

QY 11 EKRTMWEIFSETKQVFQKDL-EKIA-----PKEKGITAMSVKEVLQSLVDGMDVDCER 64
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 639 DEREKQEEIDTYKD---LKDREKVSLLQGLDSEKASLLDIKEHASSLASSGLKKDSR 695
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 65 IGTSNYWAFPSKALHARKHKLVEVLSQSEGSKHASLQKSIKAKIGRCETEERTRLA 124
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 696 LKTLLEI-----ALEQKKECLKMSQLK-----KAHEATLEA 727
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 125 KELSLRDREQLKAEVEKYKCDPQVVEIRQANKVAKAEANRWTD---NIPAIKSWAK 181
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 728 RASPEMSDRIQQLEREISRYKDESSKAQTEVDRLLLEILKEVENEKNDKDKKIAELESLS 787
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 182 RKFGFEENKI 191
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 788 RQVKDQNKV 797
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 19
Q877I1 PRELIMINARY; PRT; 1177 AA.
AC Q877I1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SMC protein.
GN SMC.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RA Cobbe N., Heck M.M.S.;
RT "Phylogenetic analysis of SMC proteins."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ543649; CAD66602.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0007059; P:chromosome segregation; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR003900; KID_repeat.
DR InterPro; IPR003345; M_repeat.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF02524; KID; 8.
DR Pfam; PF02370; M; 5.
DR Pfam; PF02483; SMC_N; 1.
DR Pfam; PF02463; SMC_C; 1.
DR TIGRFAMs; TIGR00650; MG442; 1.
SQ SEQUENCE 1177 AA; 135013 MW; EC12C4C060ESA1CB CRC64;

Query Match 11.3%; Score 118.5; DB 1; Length 1177;

```

RESULT 22
09H2G7


```

ID Q9H2G7 PRELIMINARY; PRT; 795 AA.
AC Q9H2G7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CTCL tumor antigen se2-1 (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21143360; PubMed=11149944;
RA Eichmuller S., Usener D., Dummer R., Stein A., Thiel D.,
RA Schadenroff D.;
RT "Serological detection of cutaneous T-cell lymphoma-associated
RT antigens.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:629-634 (2001).
DR EMBL: AF273043; MAG34903.1; -.
DR InterPro: IPR008827; SCP-1.
DR Pfam: PF05483; SCP-1; 1.
FT NON_TER 795
FT SEQUENCE 795 AA; 93500 MW; 45472A86A5A7ADFA CRC64;

Query Match 11.2%; Score 117; DB 4; Length 795;
Best Local Similarity 23.9%; Pred. No. 5.3;
Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13;

QY 3 KKKLSAE--KRTMMEI--FSETKDV-----FQKDLKLEKIAPKKEGITAMSVKYL 51
Db 413 QKSSSELEMTKLTNNKEVELEELKVLGKETLLYENQFKIABELKG-TEQLIGLL 471
QY 52 QSL---VDGWDCEIGTSNYWAPPSKAL-----HARKKLEVL--- 90
Db 472 QAREKEVHDLQLTATTSSEQYYSKEVKDLKTELENEKLNKNTLTSHCNKLSLENKLT 531
QY 91 -----SOLSEGSQKASLOKSIKAKIGRCETEERTERLAKELSSLRDQREQ 136
Db 532 QETSDMTLKNQOEINNNKQOEERMLQIE--NLQETETQLRNEVLEVRBELKQKDE 589
QY 137 LKAEVKYKD-CD-----PQVVEIRQANKVAKE--AANRWTDNIFAIRK----- 177
Db 590 VKCKLDKSEENCLNRKQVENKNKYIEELQENKALKKKGTAEKQLNVEIKVKNKLE 649
QY 178 -SWAKSKFG-----FEENKI 191
Db 650 LSAKQKFEITDTYQKEIEDKKI 673

RESULT 23
Q8CIY9 PRELIMINARY; PRT; 948 AA.
AC Q8CIY9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ERC1b.
GN ERC1b.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22295041; PubMed=122391317;
RA Wang Y., Liu X., Biederer T., Sudhof T.C.;
RT "A family of RIM-binding proteins regulated by alternative splicing:
RT Implications for the genesis of synaptic active zones.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14464-14469 (2002).
DR EMBL: AF541926; AAN39293.1; -.
DR InterPro: IPR002017; Spectrin.
FT SEQUENCE 948 AA; 108750 MW; A1C8DBA9FBCF2F38 CRC64;

Query Match 11.2%; Score 117; DB 11; Length 948;
Best Local Similarity 24.2%; Pred. No. 6.4;
Matches 46; Conservative 32; Mismatches 72; Indels 40; Gaps 6;

QY 11 EKTRMMEIFSETKDVLFQKDL-EKIA-----PKEKGITAMSVKVLQSLVDDGMVDCER 64
Db 611 DEREKQEIPTYKKD--LKDLKEKVSLLQGLDSEKEASLLDLKEHASSLASSGLKKDSR 667
QY 65 IGTSNYWAPPSKALHARKKLEVLSEQLSEGSQKASLOKSIKAKIGRCETEERTLA 124
Db 668 LKTLLEI-----ALEQKKEELKWSOLK-----KAHEATLEA 699
QY 125 KELSRLDRQEROLKAEVEKVKCDPQVVEIRQANKVAKEAANRWTD--NIFAIRKSWAK 181
Db 700 RASPEMDRIQQLEREIARYKDESSKAQTEVDRLLLEILKEVEENEKNDKDKKIAELESUTS 759
QY 182 RKGFEEENKI 191
Db 760 RQVKDQNKV 769

RESULT 24
Q81IU3 PRELIMINARY; PRT; 948 AA.
AC Q81IU3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CAST2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Ohtsuka T., Inoue E., Rikitsu-Takao E., Deguchi-Tawarada M., Inoue M.,
RA Takai Y.;
RT "CAST2: Identification and characterization of a protein structurally
RT related CAST.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY174115; AAO25554.1; -.
DR InterPro: IPR02017; Spectrin.
FT SEQUENCE 948 AA; 108817 MW; 7CC15D7B1939561C CRC64;

Query Match 11.2%; Score 117; DB 11; Length 948;
Best Local Similarity 24.2%; Pred. No. 6.4;
Matches 46; Conservative 32; Mismatches 72; Indels 40; Gaps 6;

QY 11 EKTRMMEIFSETKDVLFQKDL-EKIA-----PKEKGITAMSVKVLQSLVDDGMVDCER 64
Db 611 DEREKQEIPTYKKD--LKDLKEKVSLLQGLDSEKEASLLDLKEHASSLASSGLKKDSR 667
QY 65 IGTSNYWAPPSKALHARKKLEVLSEQLSEGSQKASLOKSIKAKIGRCETEERTLA 124
Db 668 LKTLLEI-----ALEQKKEELKWSOLK-----KAHEATLEA 699
QY 125 KELSRLDRQEROLKAEVEKVKCDPQVVEIRQANKVAKEAANRWTD--NIFAIRKSWAK 181
Db 700 RASPEMDRIQQLEREIARYKDESSKAQTEVDRLLLEILKEVEENEKNDKDKKIAELESUTS 759
QY 182 RKGFEEENKI 191
Db 760 RQVKDQNKV 769

RESULT 25
Q84HV4 PRELIMINARY; PRT; 920 AA.
AC Q84HV4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
```



```

RP SEQUENCE FROM N.A.
RA Toramoto T., Ikeda D., Ochiai Y., Minoshima S., Shimizu N., Watabe S.;
RT Multiple genes organization of Torafugu Fugu rubripes tropomyosin and
RL tissue distribution of their transcripts";
DR EMBL; AB090362; BAC57564.1; -
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00261; Tropomyosin; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PROSITE; PS00326; TROPOMYOSIN; 1.
SQ SEQUENCE 284 AA; 32997 MW; DA08CDD308782C7 CRC64;

Query Match 11.0%; Score 115; DB 13; Length 284;
Best Local Similarity 23.3%; Pred. No. 2.3;
Matches 51; Conservative 42; Mismatches 86; Indels 40; Gaps 6;

QY 3 KKGLSAEEKTRMWEIFSETKDVFLKDLKIAPEKKGITAMSVKVEQLSLVDDGMVDC 62
DB 47 EKRLVTEEDRDVTEEFQTAES--KLTAEVATRAEAVA-SLNRRLQVVEELDRAQ 103

QY 63 ERIGTSNYWAFPSKALHARKHKLVEQLSEGSQKSHASLQKSIKAK----- 111
DB 104 ERLGALTLEBAEKAADSESGMKVIENRAMKDEEKVLEQLKEAKGHIAEEADRYKE 163

QY 112 -IGRC-----EFEERTRA-----KELSSLRQREQLKAEVEKYKDCDPQVVE 153
DB 164 EVARKLVIIIGDLERTEERAELASCAELEELKVTNNLKSLEAQAESQKEDKYEE 223

QY 154 EIQQANKVAEAAANRWTDNIFAISWAKRGFGFEENKID 192
DB 224 EIKVLTDKLEAEETR-----AFAERTVAKLEKID 254

RESULT 29
Q96N90 PRELIMINARY; PRT; 365 AA.
ID Q96N90
AC Q96N90
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ31229 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055791; BAB71015.1; -
KW Hypothetical protein.
FT NON TER 365
SQ SEQUENCE 365 AA; 42269 MW; 42AE6F3A2A679730 CRC64;

```

```

Query Match 10.9%; Score 114.5; DB 4; Length 365;
Best Local Similarity 21.8%; Pred. No. 3.3;
Matches 44; Conservative 42; Mismatches 63; Indels 53; Gaps 8;

QY 11 EKTRMWEIFSETKDVFLKDLKLE---KIAPKEKGITAMSVKVEQLSLVDDGMVDCERIG 66
DB 58 ESKSKLEKESLVQKEGYEVESSLLKDAEKEATE-----AQSLVENQMATCEKLN 111

QY 67 TSNYWAFPSKALHARKHKLVEQLSEGSQKSHASLQKSIKAKIGRCETERTLAKE 126

```

```

DB 112 RSN-----SELEDEILCLEKELKEEKSHSE-----QDELMAIDSKR 148
QY 127 LSSLRDQREQLKAEV-----EKYKDCDPQVVEE---IROANKVAEAAANRW 169
DB 149 IQSLEDSKSLKSQVAEAKMTFKIFQNWEEKLKIAIDALNENSQLESQKQLQAEVW 208
QY 170 TDNIFAIKSWAKRGFGFEENKI 191
DB 209 KEQVSELN---KQKVTFEESKV 227

RESULT 30
Q84H67 PRELIMINARY; PRT; 448 AA.
ID Q84H67
AC Q84H67
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Major surface protein 3 (Fragment).
GN MSP3.
OS Anaplasma marginale.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Florida;
RX MEDLINE=22423820; PubMed=12535066;
RA Meesus P.F., Brayton K.A., Palmer G.H., Barbet A.F.;
RT "Conservation of a gene conversion mechanism in two distantly related
RL Mol. Microbiol. 47:633-643(2003).
DR EMBL; AY127888; AAM97265.1; -.
FT NON TER 448
SQ SEQUENCE 448 AA; 51698 MW; A76DCBE33030AES CRC64;

Query Match 10.9%; Score 114.5; DB 2; Length 448;
Best Local Similarity 25.6%; Pred. No. 4.2;
Matches 43; Conservative 38; Mismatches 64; Indels 23; Gaps 6;

QY 3 KKGLSAEEKTRMWEIFSETKDVFLKDLKIAPEK-KGITAMSVKVEQLSLVDDGMVD 61
DB 51 KQKTTKAKEKVKRELKEKIEEEQKWLKLBQLKPEETKXKLGSEIEKLRALVKEGLGD 110
QY 62 CERIGTSNYWAFPSKALHARKHKLVEQLSEGSQKSHASLQKSIKAKIGRCETERT 121
DB 111 LERL-----EAKKLEEVKVKVEVAIKVEDIEE--VDDQVGELEKLEES 152
QY 122 RLAKELSLRD-QREQL--KAEVEKYKDCDPQVVEEIROANKVAEAA 166
DB 153 ELLKKIAIGDLEQEQLEMEKEIEKL-ESGTQLQOERMKMLTERRKKLA 199

Search completed: September 27, 2004, 08:37:46
Job time : 48 secs

```

